

Time to grasp the international perspective on GM crops

Among the clamour of voices from around the globe for and against GM crops, many scientific academies have been quiet. Their views are urgently needed in determining the long-term research agenda.

Next month sees an event that should have happened years ago: scientists representing national academies from developed and developing countries are to meet to discuss the way forward for global agricultural biotechnology. The agenda includes discussions on the technology's promises, its potential risks and its regulation. Participants will consider adopting a common position on those issues and the scientific uncertainties in genetically modified (GM) crops. They may even agree to coordinate an international study, for example into the environmental implications of genetic modification in agriculture.

The importance of this meeting — and of such a study — can hardly be overstated. It is commonly argued that GM crops are integral to the future of world food production and have a major role to play in feeding the hungry. But the institutional voices carrying this message — such as Britain's Royal Society, and more recently the Nuffield Council on Bioethics — have come largely from the developed world.

Some influential non-government organizations — such as the Third World Network and the Research Foundation for Science, Technology and Natural Resource Policy in India — have disputed the vision that GM crops are needed to alleviate hunger and malnutrition. Those organizations, like their counterparts in the developed world, claim that hunger is caused because the poor are unable to get access to food, and not necessarily because of a shortage of production.

While many prominent individual scientists have made their views clear, the perspectives of the developing world's leading science

academies have been largely absent. Next month's meeting should help to correct that anomaly.

But there is arguably a second, and equally important, reason why next month's meeting could not be more timely: the United States is threatening Europe with — another — trade war unless states of the European Union relax regulations on the sale of GM crops, which are stricter than those on the opposite side of the Atlantic.

European countries cite broad environmental concerns, and last week heads of state of Europe's G8 countries tried to persuade the United States to set up an international committee of scientists to validate new foods (see page 717). The US government, on the other hand, believes Europe's environmental concerns are little more than a smokescreen for old-style protectionism.

Neither side is able to draw on a truly global scientific assessment of potential environmental risks from GM crops because, as yet, none exists. Indeed, the absence of a credible international scientific assessment on the issue contributed to the collapse of biosafety talks earlier this year. Europe, Africa and the United States disagreed on even the most fundamental question: whether GM organisms contribute to global biodiversity loss.

Countries will resume discussions on the biosafety protocol in May 2000. But without such a study, it is difficult to see the meeting achieving a positive result. An international scientific review of the environmental impacts of GM crops is urgently needed, and science academy representatives meeting next month have the opportunity to take the initiative. They should set the ball rolling without delay. □

Confrontation out, pragmatism in

A year-long 'stand-off' between French researchers and the science ministry signals the need for change.

Just three weeks ago, the French science ministry created a special FF40 million (US\$6.3 million) fund for grants to young researchers with original ideas in any field. It has been taken aback by the scale of the demand — some 1,000 proposals have already been received. This small injection of oxygen has clearly set alight many young minds who are otherwise too often stifled by a patriarchal French laboratory system where the grip of powerful laboratory directors on the strings of the purse and the ideas suffocates innovation.

There is also a message here for a country where everyone is growing tired of a repetitive cycle where the government introduces sweeping reforms, researchers take to the street to oppose them, and deadlock sets in until the next government arrives and the ritual begins over again. That message is that pragmatic and well-targeted reforms may be a more effective way forward than grandiose rearrangements of the research administration.

A similar conclusion seems to be emerging from a national consultation, ordered by Lionel Jospin, the prime minister, in a bid to find a way out of the current deadlock over proposed reforms by Claude Allègre, the science minister. Organized by members of par-

liament Pierre Cohen and Jean-Yves Le Déaut — who are also working scientists — the consultation will culminate this weekend with a national colloquium in Paris. The take-home message is likely to be that modernizations of practices, each tuned to particular situations at the grassroots, are needed more than high-profile reforms of, for example, whole funding agencies.

Ironically, the problems acknowledged as priorities by researchers are virtually identical to those that most trouble Allègre's ministry. The lack of independence of young scientists is one. The bloated system of evaluation is another. And then there is the long-recognized rigidity of a system where some scientists enjoy full-time posts in the research agencies while their colleagues in the universities — often young scientists in their prime — struggle to fit research in alongside excessive teaching loads.

Valuable months have been lost as a result of a sterile confrontation between researchers and the ministry. For the sake of science and of government credibility, this weekend's meeting must herald a more productive period where the ministry and the research community can identify critical areas where tangible progress can be achieved. □



Gore under fire in controversy over South Africa AIDS drug law

[WASHINGTON] US Vice-President Al Gore's campaign for the presidency has become caught up in an escalating controversy over his opposition to a South African law aimed at providing low-cost AIDS drugs.

Gore, who officially launched his campaign last week, has actively opposed South Africa's Medicines and Related Substances Control Amendment Act. This seeks to bypass existing patents to allow the manufacture or import of expensive AIDS drugs at significantly lower costs than those currently charged by the pharmaceutical industry.

The law was approved by the country's parliament in 1997. But it has not been implemented because 41 pharmaceutical companies, led by the Pharmaceutical Manufacturers' Association of South Africa (PMASA), have challenged it in the High Court in Pretoria.

The law would allow compulsory licensing and parallel importation of medicines by South Africa. This means either that manufacturers inside the country could bypass patents and make the drugs more cheaply, paying only royalties to the patent holder; or that South Africa could import the drugs from third countries in which they are already less expensive.

The AIDS epidemic has exploded in South Africa, where recent surveys in pregnancy clinics indicate that 22 per cent of sexually active adults are infected (the figure was 14 per cent in 1997). But the latest AIDS therapy costs US\$800 a month in South Africa, where the average annual income is \$2,600.

"We have a Third World country that has this horrifying pandemic, and we can't afford the drugs," says Robert Shell, a



Campaign chaos: Gore's bid for the presidency got off to a bad start last week. Protesters disrupted his candidacy announcement, waving placards condemning his opposition to a South African law that could let AIDS drugs be sold more cheaply.

demographer at Rhodes University in Grahamstown, South Africa, who tracks the epidemic. "It is a question of life and death."

But the South African law has riled the US government and the pharmaceutical industry. They claim that it is striking at the heart of the patent protection necessary for drug development. Led by Gore and the US trade representative, the US has pressed South African officials to rescind or rewrite the law.

A spokesman for Gore says he and Thabo Mbeki, the South African president, "are committed to working together to chart a course that will meet the medical needs of those infected with HIV or AIDS, without cutting off the commercial incentives that fuel medical research".

Gore and Mbeki co-chair the US-South Africa Binational Commission, an economic diplomacy group. The spokesman adds that Gore "has a very strong record of efforts to fight AIDS in South Africa and his stands have been consistent with that".

But the perception that Gore is leading an

effort to deny cheap drugs to a country in a desperate plight is making its mark in the United States. International health and AIDS activists dogged Gore on the campaign trail last week, chanting "Gore's greed kills".

"History will judge people harshly as to how they acted in this crisis. And it's going to be a harsh judgement on Gore," says James Love, director of the Consumer Project on Technology, an advocacy organization affiliated to Ralph Nader.

Love and Nader wrote to Gore in April accusing him of using "an astonishing array of bullying tactics" to stop South Africa expanding access to AIDS drugs.

Even members of the Presidential Advisory Council on HIV/AIDS are challenging Gore. "He is wrong," said council member Debbie Runions, a former Gore campaign volunteer, speculating whether "he's in the pockets of the pharmaceutical companies".

Critics allege that the vice-president, who has typically embraced liberal causes such as the fight against AIDS, is being swayed by ties to the drug industry. For instance, a key lobbyist for the Pharmaceutical Research and Manufacturers of America (Pharma), Tony Podesta, is the brother of the president's chief of staff, John Podesta, and is also a Gore adviser and friend.

According to the Center for Responsive Politics, a Washington-based public-interest group, Gore's political action committee received \$56,000 from individuals connected to pharmaceutical companies in 1998.

The vice-president's office denies that he has been influenced by drug makers. Officials point out that when Pharma asked the government last year to threaten South Africa with trade sanctions, it was rebuffed.

The country was merely placed on a 'watch list' of countries suspected of violating US intellectual property rights — a category that does not trigger sanctions. But two months later the White House announced

G8 leaders seek study on effects of biotech

[LONDON] A proposal from the French president Jacques Chirac to set up an international council of scientists to validate new foods was rejected at last weekend's summit of leaders of the G8 industrialized countries in Cologne.

But the participants pledged a "science-based, rules-based" approach to addressing the consequences of biotechnology. The meeting also agreed to commission a study of the implications of biotechnology from the Organization for Economic Cooperation and Development. The findings will be considered at the next summit.

The Chirac proposal fell foul of the US delegation, which described it as protectionist. But environmental issues

were prominent in the meeting's final communiqué.

The meeting pledged to clarify the relationship between the rules of the World Trade Organization (WTO) and United Nations environment conventions. This is a reference to the impasse between the Trade Related Agreement on Intellectual Property Rights of the WTO, and the UN biodiversity convention over the sharing of benefits from patents on indigenous plants. The meeting also pledged "early entry into force" of the Kyoto Protocol on climate change.

Education was another major theme. Delegates promised to promote more teacher and student exchanges between the G8 and other countries. **Ehsan Masood**

that it was denying preferential tariff treatment for four South African imports.

South Africa's new health minister, Manto Tshabalala-Msimang, told *Nature* that the law is meant to address the "legacy of apartheid". "We don't want to interfere with the patent rights, but we want to get drugs of quality at a price we can afford," she says.

"Why would you want to buy the drug at SAR80 (US\$13) when if you import it you would get it at SAR30?" she asks. "In situations of emergency, what do you do? This is the question I would pose to PMASA."

Nonetheless, she says, she is making it "very high on my priority list" to meet the drug makers. She says the government is willing to negotiate the implementation of the law "in such a way that it meets the needs of the pharmaceutical industry".

That will take some doing, according to the drug companies. "This is a major issue," says Jeff Trewihitta, spokesman for Pharma. "The cost of research is very high. Patent protection is the lifeblood of the industry."

Mirryena Deeb, chief executive officer of PMASA, says the law is so broad that it "undermines the patent law system" and allows the health minister to "just take away the patent right".

An official at the office of the US Trade Representative (USTR), which has aggressively lobbied South Africa to reverse the law, calls it "offensive". "They've decided that the way to solve the [AIDS] problem is to deny intellectual property rights."

The USTR, Gore and the drug groups argue that the law breaches an important World Trade Organization (WTO) agreement on trade-related aspects of intellectual property protection, the 'TRIPS' agreement.

But Gore's critics read the agreement as allowing precisely the kind of action South Africa has taken. The agreement says compulsory licensing is permitted in cases of "national emergency or extreme urgency".

Some observers say that the TRIPS language is sufficiently ambiguous for there to be honest disagreement on whether South Africa is in breach of it. What angers some, says one international health official, is that "commercial interests are leading [the US government to use] pressure tactics", rather than using WTO dispute mechanisms.

According to a State Department report presented to Congress in February, Gore, the USTR, the Department of Commerce and the State Department have mounted an "assiduous, concerted campaign" to overturn the South African law.

Some argue that the issue goes beyond patent rights. Tom Coates, director of the AIDS Research Institute at the University of California, San Francisco, warns of the danger of drug resistant viral strains developing if the drugs are not imported in the right combinations and quantities, and administered with proper oversight.

Meredith Wadman

Chlorine industry says EPA rules ignore good science

[WASHINGTON] The US Environmental Protection Agency (EPA) is being taken to court by the Chlorine Chemistry Council, a Washington-based group representing chlorine manufacturers and users. The case could determine not only how the agency assesses cancer risks, but how it weighs scientific evidence in crafting regulations.

A federal court will hear arguments next year charging that the EPA failed to use the best available scientific information when it set goals for the maximum amount of chloroform allowed in drinking water.

Chloroform, the most common chemical by-product of water disinfection, causes cancer in laboratory animals when administered at very high doses. The scientific issue concerns whether there is a dose below which chloroform is safe.

EPA guidelines for assessing cancer risk extrapolate linearly the harmful effects seen at high doses, so that any dose is considered harmful. But a non-zero safe 'threshold' could theoretically be set if enough were known about a chemical's mode of action.

The controversy over chloroform began in March 1998, when the agency released new data on disinfection by-products, and announced that it was considering changing the goal for chloroform contamination in drinking water from zero to 300 milligrams per litre — the first acknowledgement of a threshold dose for a regulated carcinogen.

But in last December's final ruling, the agency retreated from this position, and set the goal at zero. The Chlorine Chemistry Council, water utilities and other groups claimed the agency had buckled to pressure from environmentalists, and filed suit. Complying with the new regulations is estimated to cost industry \$1 billion a year.

The International Society of Regulatory Toxicology and Pharmacology (IS RTP), which, along with the chlorine council, sponsored a meeting in Washington on the EPA's chloroform policy, said in a statement that there was "no justification for the EPA management's decision to ignore the advice of its own scientists".

Agency scientists conducted an extensive review of chloroform in the five years leading up to the March announcement, considering more than 30 toxicological studies and co-sponsoring a 1997 review by the International Life Sciences Institute. Based on that analysis, the EPA concluded that "a nonlinear [threshold] approach is more appropriate for extrapolating low-dose cancer risk than the low-dose linear approach".

Agency scientists involved in the work say they were frustrated, therefore, when the



Bottling out: as sales of bottled water rise, court case focuses attention on the safety of tap water.

EPA reverted to the linear approach in its final ruling.

Erik Olson of the Natural Resources Defense Council (NRDC), a Washington-based environmental group that opposes a non-zero goal, says the episode reveals "deep divisions" within the agency.

Language in the December ruling bears this out. Calling the zero goal an "interim risk-management decision," the agency said that even though the "underlying science for using a nonlinear extrapolation approach... is well founded," further discussion was needed, as the new approach would represent a "significant and precedential, albeit sound, application of new science to the policy and risk-management decision-making process".

The Chlorine Chemistry Council hopes to exploit that candidness, arguing that the agency has not used the best available peer-reviewed science, as laws governing drinking water safety require. Oral arguments will be heard in a federal circuit court in Washington, which recently overturned EPA regulations on air pollution, partly on the basis of inadequate justification.

Olson believes the EPA made the right choice in December, and points out that not all scientists agree that the mode of action for chloroform carcinogenicity is well enough established to end the zero-tolerance policy.

But many toxicologists see the EPA's action as a stalling tactic. "When we have data, the simple use of a default represents a failure," said Jay Goodman, president of the Society of Toxicology.

EPA guidelines for assessing cancer risk have been under review for nine years, and have been further delayed by an additional review to consider children's health. The agency's science advisory board last month urged the EPA to publish the guidelines "as soon as judiciously possible". **Tony Reichhardt**

Wellcome Trust boosts researchers' pay

[LONDON] In a move widely seen as a direct challenge to the UK government to increase researchers' salaries, Britain's largest medical research charity, announced last week that it is to increase the salaries of some of its research staff to 30 per cent above the basic university pay scale.

The announcement comes shortly before the delivery of a major report on UK pay and conditions in higher education, and during a pay dispute between employers and members of the higher-education union, the Association of University Teachers (AUT). The Independent Review of Higher Education Pay and Conditions, known as the Bett report after chairman Sir Michael Bett, is expected to recommend increases of up to 20 per cent for some academic staff.

The Wellcome Trust's pay rise will cost around £6 million (\$9.5 million) at first and then an extra £2.5 million a year. It will benefit 270 of the trust's 3,000 or so researchers who are in its Career Development Scheme (CDS) and on contracts of three years or more. The trust says it will bring salaries of CDS researchers into line with other professions. Salaries in British universities are estimated to be 30 to 40 per cent lower than their professional counterparts outside the academic community.

"Science has become a very unattractive



career," says Michael Dexter, the trust's director. "When we found out that some of our PhD students were taking a drop in pay [when getting academic appointments] we knew something was very seriously wrong."

Dexter says the trust has had recruitment problems, particularly when recruiting overseas researchers. "We are simply fetching back wages to levels 15 to 20 years ago," he says.

"Academic and research staff are not paid enough in comparison with other professional classes," says Peter Cotgreave, director of the pressure group Save British Science. "Since 1989, departments have been having trouble filling posts with good candidates."

The Bett committee was set up last year to examine pay levels, pay structures and conditions of service. Its report, due to be delivered yesterday (23 June), is expected to warn the government that a failure to find additional money will mean that institutions will struggle to maintain the quality of research and teaching. Nevertheless, even if government were to find additional money, much of it could go on equalizing pay between men and women in higher education — a gap which has been noted by leaked versions of the report.

When the report comes out this week, "we will look at expanding [the increase] to include programme and project grant applicants," says Dexter. He concedes that a wider pay increase could mean fewer, better-paid researchers. He adds that, if the best researchers are lost to other areas and overseas, then "we'll end up in a pretty poor state".

The AUT has congratulated the Wellcome Trust on the increase, and has rejected an offer of a 3.5 per cent increase for academic and related staff in higher education. AUT members within pre-1992 universities are demanding a 10 per cent increase, and a second round of action is planned for August. "This is a crucial time for higher education," says David Triesman, general secretary of the AUT. "The rot of depressed pay has to stop now."

Natasha Loder

Japan tightens rules on GM crops to protect the environment

[TOKYO] Japan is to tighten its safety regulations on genetically modified crops following the publication last month of research suggesting that pollen from *Bt* corn could harm the larvae of monarch butterflies (see *Nature* 399, 214; 1999).

The Japanese Ministry of Agriculture, Forestry and Fisheries (MAFF) announced last week that it will suspend approval of *Bt* crops for agricultural purposes until its committee on genetically modified organisms (GMOs) has established criteria for evaluating the safety of such crops.

Japan has already approved the importation of six types of *Bt* corn for use as foodstuffs, but the commercial planting of seed produced by US companies, such as Monsanto, has not yet been approved.

Yutaka Tabei of the ministry's safety evaluation division says the harmful effect of *Bt* toxins on non-target insects was not entirely unexpected. "The results were not surprising, given that the butterfly larvae were fed leaves dusted with pollen from *Bt* corn," he says. "But we must carry out further studies — including those on the spread of pollen — to assess any potential impact such crops may have in the natural environment."

There is a strong emphasis on the concept of 'substantial equivalence', under which GM foods are compared with analogous conventional foods in terms of characteristics such as toxicity and nutritional qualities. At present, the GMO committee predicts the potential ecological impact of *Bt* crops to be "negligible" in the natural environment. But it is emphasizing the importance of carrying out safety tests by simulating various environmental conditions.

The committee plans to release revised safety evaluation protocols by the end of this year, basing its final decision on safety studies carried out by Japanese institutions.

The move represents the first major step by the government to review the potential ecological risks of GM crops. Until the launch of a research project in April to examine the long-term effects of herbicide- and insect-tolerant crops on ecology and agricultural practices (see *Nature* 398, 655; 1999), the main safety concern about GM foods had focused on the risk to health.

Debates about GMOs had therefore centred on the labelling of products that contain genetically modified ingredients. Such foods are currently not labelled in Japan, and MAFF is expected to decide by

the end of the year whether to require products containing GMOs to be labelled as such (see *Nature* 395, 628; 1998).

Japan's regulations on GMOs, which are overseen by the Ministry of Health and Welfare for food safety and by MAFF for field use, are based on guidelines set out by the Organization for Economic Cooperation and Development (OECD).

Although MAFF requires farm-scale trials of GM crops, critics say these are inadequate, as its safety evaluation protocols overlook proper assessments for long-term and 'unexpected' risks to genetic diversity.

"Although MAFF's safety evaluation of *Bt* corn requires tests on its impact on non-target species such as mice and ladybirds, it excludes tests on butterflies by ruling out the possibility of pollen depositing on other plant species," says Setsuko Yasuda, director-general of Japan's Consumers' Association.

Many see MAFF's decision to review its safety protocols as a step towards gaining public support for developments of GM technology in Japan. Japan Tobacco are planning to develop GM rice, and other companies are embarking on research into GM trees for high pulp yield and insect-tolerant GM flowers.

Asako Saegusa

Urgent talks follow damning report on US weapons labs

[WASHINGTON] The US Department of Energy and congressional allies of its troubled nuclear weapons laboratories are this week seeking to hammer out a compromise for their future administration. They are keen to deflect growing pressure for drastic change, following the release of a scathing report on security at the laboratories.

The report, *Science at its Best, Security at its Worst*, was issued last week by the President's Foreign Intelligence Advisory Board (PFIAB), chaired by a former Republican senator, Warren Rudman. The report "struck quite a nerve" in the administration, according to one official there, by the intensity of its criticism of the energy department's management of the laboratories.

Even more troubling for the administration was the board's recommendation that the labs should be split off from the main structure of the department and placed in a semi-autonomous agency within the department.

Lab supporters had earlier blocked a drive by Republicans to move all nuclear weapons research into the defence department and end visits to the labs by foreign scientists.

But the Rudman report has revived pressure for administrative reforms. Senator Pete Domenici (Republican, New Mexico), a leading supporter of the labs, has come out in favour of Rudman's proposal for an Agency for Nuclear Stewardship within the energy department. It would be similar in concept to the National Oceanographic and Atmospheric Agency within the Department of Commerce.

At first, Bill Richardson, the energy secretary, called the plan unworkable. "The security problems at the department are broader than the board recognizes," he said. A new agency would "erode the link between national security and science at its best".

But by last weekend Richardson was said to be backing down and a compromise with Domenici was in the works. One option was to create an undersecretary for weapons and security in the energy department. But such a move could be construed as adding another layer of the administrative complexity which Rudman roundly criticized. The issue was due to be addressed this week at joint hearings of four congressional committees.

Meanwhile, developments unfolded in the investigation of security violations at the Los Alamos National Laboratory, New Mexico. The *Albuquerque Journal* reported that the inquiry was extending considerably beyond Wen Ho Lee, the scientist whose alleged security violations triggered the lab scandal. On Monday (21 June), Richardson said that polygraph testing was to begin this week on up to 5,000 weapons lab staff.

Will Lepkowski

Varmus defends plan for global biomedical e-journal

[WASHINGTON] Harold Varmus, director of the US National Institutes of Health (NIH), has responded to critics of his plan for a global website centralizing biomedical literature, arguing for example that it would not threaten traditional journals.

In a seven-page message posted on the NIH's website on Monday (21 June), Varmus says that 'E-Biomed' would give "free, fast and full access" to the entire biomedical research literature for anyone with a computer and an Internet connection.

And, while accepting that the cost and mechanism of paying for the site are undetermined, he suggests that financing could come from fees levied on authors.

Varmus's retort to his critics follows the NIH's receipt of more than 200 responses to the proposal (see *Nature* 398, 735; 1999 & 399, 8-9; 1999). The NIH director wants to create an "electronic public library of medicine" containing peer-reviewed and non-peer-reviewed literature.

Many journals see the proposal as threatening their existence. But Varmus stresses that 'E-Biomed' should "in no sense be interpreted as a proposal to interfere with, control or restrict journals". The NIH is eager to encourage journals, especially top-tier ones, "to become part of the ['E-Biomed'] system".

He also calls concerns that 'E-Biomed' represents a power grab by the federal government an "unfortunate misreading". The site, he writes, "would not be owned by the NIH or any other component of the US government". The agency's role would be limited to "technical assistance and financial support".

Varmus challenges the worries of scientific societies that they would be driven into

poverty by the loss of subscriptions, counselling them to raise membership fees and conduct workshops to generate new income.

And he downplays concerns about the non-peer-reviewed section of 'E-Biomed'. Critics have said that unreviewed clinical papers could endanger public safety, and that unreviewed work would erode the line between rigorous and 'junk' science. Few scientists would publish the latter, says Varmus, because of concern for their reputations. And clinical worries are "misplaced" because readers will be clearly informed about which reports have been peer-reviewed.

But Varmus's response drew a negative reception from several critics. "It's a pipe dream to think I can double dues" to recover lost subscription income, says Martin Frank, executive director of the American Physiological Society, which publishes 14 journals.

Jerome Kassirer, editor-in-chief of the *New England Journal of Medicine*, is still concerned about the non-peer-reviewed section. "It contaminates the literature," he says, adding that, in the case of clinical studies, it potentially endangers the public, whom he does not believe are adept at discriminating between dubious material and solid science.

But Michael Cox, a professor of biochemistry at the University of Wisconsin-Madison, who wrote to Varmus last month calling 'E-Biomed' "among the worst ideas I have ever heard," says he is somewhat mollified by the NIH director's retort. "The people putting this together are thinking about [scientists' and journals'] major concerns," says Cox.

Varmus's statement can be found on the web at www.nih.gov/welcome/director/ebiomed/ebiomed.htm

Meredith Wadman

One-stop shop for 200 life science journals

[PARIS] The American Institute of Biological Sciences has announced plans to combine up to 200 of the journals published by its member societies within a single, searchable website. A prototype version of 'BioOne', as it is called, is scheduled to be launched in 2001.

The institute's 55 members include the American Museum of Natural History and the Botanical Society of America. BioOne is supported by a consortium made up of the institute and the Scholarly Publishing and

Academic Resources Coalition (SPARC), the University of Kansas, the publisher Allen Press, and the 'Big 12 Plus Libraries Consortium', which groups 23 US research libraries.

The move is a dramatic example of a trend towards aggregation of titles on the web that is being driven by user demand for 'one-stop shopping', the capacity to search a wide literature from a handful of gateways.

BioOne also intends to digitize and make available back issues of journals, and

is keen to include journals published by non-members of the institute. It is the latest in a series of initiatives by SPARC and US universities to put pressure on the publishers of expensive journals by offering competitive web publishing services at reduced prices (*Nature* 397, 195-200; 1999).

BioOne's goal is to help learned societies and the publishers of inexpensive journals to establish a presence on the web by combining expertise and infrastructure.

Declan Butler

Compromise sought on 'Terminator' ...

[LONDON] Scientific advisers to the United Nations Biodiversity Convention have given a mixed response to a more farmer-friendly alternative to the controversial 'Terminator' technology, in which seeds are genetically modified to become sterile after one season's planting.

The alternative is being dubbed 'Trait-specific Genetic Use Restriction Technology', or T-Gurt. Like Terminator seeds — another type of Gurt — T-Gurt seeds are also being developed by the agro-chemical industry, notably the US company Monsanto and the British company AstraZeneca. Unlike Terminator seeds, however, T-Gurt seeds do not become sterile after a season's planting.

T-Gurt seeds are genetically modified to produce specific traits, such as tolerance to salt or drought. If a farmer wished to activate the trait in one type of T-Gurt, he would have to spray the seed with a proprietary chemical. The seed will still germinate without the chemical, but it would not have the modified characteristics.

Representatives of countries in the European Union, Latin America and southeast Asia gave a cautious welcome to T-Gurts at a meeting of the biodiversity convention's scientific advisory board in Montreal this week.

But scientists from African countries have joined conservation groups such as the Rural Advancement Foundation International (RAFI) in Canada in urging the convention's scientific advisers not to endorse the technology until they are satisfied that T-Gurts do not harm human health or the environment.

Critics also include members of the Third World Academy of Sciences. Members of the academy's governing council are expected to debate a proposal tomorrow (25 June) to outlaw the patenting of seeds intended to be grown as food crops.

"Agriculture in much of the developing world is the result of collective experience

gained from the sweat and toil of poor peasants over thousands of years," says Muhammad Akhtar, vice-president of the academy and emeritus professor of biochemistry at Britain's University of Southampton.

Akhtar, who will be arguing in favour of the proposal, adds that "the gene pool in present foods is a common heritage of humankind and should belong to it all".

Supporters of T-Gurts include Richard Jefferson, director of the Center for the Application of Molecular Biology to International Agriculture in Canberra. Jefferson is the main author of a report on the environmental implications of Gurt technologies published by the biodiversity convention secretariat in time for this week's meeting.

Jefferson says he sees T-Gurts as a compromise technology that meets industry's desire to protect intellectual property and maximize returns on innovations in agricultural biotechnology, without jeopardizing the practice of subsistence farmers using seeds from a crop for planting in the future.

A key attraction of all Gurts is that they give companies better protection against unauthorized copying than patents. This desire for a tighter system of intellectual-property protection is a major concern for the technology's opponents.

Patents are conferred on inventions that demonstrate novelty, an inventive step or non-obviousness. But Terminator and T-Gurt technology, at least in principle, may be applied to any seed, regardless of its novelty.

In addition, Terminator technology can be licensed indefinitely, whereas patents have a finite life of usually not more than 20 years. Similarly, Gurts provide their owners with protection from unauthorized copying, but without the costs and effort of pursuing a patent infringement.

But it is the threat to seed saving — commonplace among poor farmers worldwide — that is sustaining the pressure on the



Seeds of discontent: a Mexican protester boards a ship importing genetically modified maize.

biodiversity convention's member states to outlaw Terminator technology.

Terminator is the subject of a United Nations resolution. And last October the Consultative Group on International Agricultural Research, a network of centres in the developing world organized through the World Bank, promised not to incorporate "into its breeding materials any genetic systems designed to prevent seed germination".

Much of the pressure is coming from RAFI, which has written to ministers and officials in departments of agriculture, environment and patent offices in 140 countries urging them to "assert national sovereignty over their seed supply and to ban the seed sterilization technology outright".

"Many governments are unaware that the World Trade Organization allows countries to reject individual patents on the grounds that they are contrary to public morality and/or a threat to health or the environment," says Pat Mooney, RAFI's executive director.

Arguably the best-known Terminator patent was granted last July to the US Department of Agriculture and the seed company Delta and Pine Land — itself the subject of a takeover bid from Monsanto.

Monsanto has so far resisted calls for a moratorium on commercializing its technology. Sources say it will not do so without similar assurances from its competitors, but so far only AstraZeneca has indicated that it does not intend to use the technology to deprive farmers from replanting seeds.

Monsanto executives also do not want to risk losing investor confidence by undermining an important technology of one of their potential acquisitions. **EhsanMasood**

... as academies meet to plan a global approach

[LONDON] Representatives of the science academies of the United States, Britain, Brazil, China, India and Mexico and the Third World Academy of Sciences are to meet in London next month to discuss a possible joint study on genetic modification (GM) in world agriculture.

The meeting will be held on 13 July at the Royal Society in London, which is organizing the gathering jointly with the US National

Academy of Sciences.

Several recent studies — including ones from the Nuffield Council on Bioethics (see *Nature* 399, 396; 1999) and the Royal Society (see *Nature* 395, 5; 1998) — have concluded that GM technology has a significant role to play in tackling world hunger. But this meeting will be the first time that academies from developing countries have been invited to contribute their ideas.

Each academy has been asked to send up to three experts. The meeting will discuss the possibility of a common position on issues such as the extent to which GM crops can contribute to food production, and the environmental risks of GM crops, along with specific issues such as 'Terminator' seeds (see main story) and regulatory matters, such as the United Nations biosafety protocol. **E.M.**

Five bid to host Middle East synchrotron

[PARIS] A proposal for an international research centre in the Middle East, built around a synchrotron to be donated by Germany, was officially launched at a meeting in Paris last week. Bids to host the centre were submitted by Turkey, Cyprus, Iran, the Palestinian Authority and Egypt.

A final decision on the location will be taken by an interim governing council, including two representatives from each of the candidate countries (see *Nature* 399, 507–508; 1999).

Delegates have promised to support the project even if the site does not fall in their home country. In a bid to accelerate the approval process, the meeting called on states in the Middle East to confirm their

approval by 31 July. There was considerable praise for Israel's commitment to the project without bidding to host the facility, out of recognition that a centre in Israel would attract little participation from many of its neighbours.

"The project would not work without Israel," says Siegbert Raither of the United Nations Educational, Scientific and Cultural Organization (Unesco), which hosted last week's meeting. "Israel is a very important ingredient to the scientific success of this institution," he says.

Participants at the meeting acknowledged that funding is now their biggest priority. But raising the required amount — at least US\$30 million — will not be easy.



Schopper: keen for a quick decision.

Funds for the project are unlikely to emerge until a decision has been made on the centre's location. This is partly because Western governments may be reluctant to commit money to a centre in any country where diplomatic ties are weak.

The proposal is based around an offer from Germany to donate BESSY-1, a 14-year-old, fully functioning, 0.8 GeV synchrotron in Berlin. But Germany may not be prepared to commit the extra costs of dismantling the machine for later use without a firm indication that the centre will be funded.

"The BESSY directorate is looking for a quick decision, as there is a [financial] difference between the machine being scrapped or being dismantled for further use," says Herwig Schopper, a former director-general of the European Laboratory for Particle Physics (CERN). Schopper, who was elected president of the project's interim council last week, emphasizes that the timetable is tight. BESSY-1 is due to be taken out of service at the end of the year.

Potential sources of funding include the European Union and the Middle East finance package tied to the Wye Agreement, which is working its way through the US Congress. President Bill Clinton is understood to have requested \$1.2 billion for Israel, \$400 million for the Palestinian Authority and \$300 million for Jordan.

Observers believe that the costs of the war in Kosovo could mean that Congress is unlikely to agree to the sum requested by Clinton. And most of the Wye package is understood to be earmarked for roads, hospitals and schools. William McIlhenny, the US government's permanent observer to Unesco, says a key question is whether governments are prepared to allocate large sums for a synchrotron facility instead of more pressing problems such as clean water and sewage disposal.

A centre based in the Palestinian Authority is likely to be convenient for Israel. But the Palestinian bid, along with those from other candidate hosts — except Egypt, and possibly Iran — will need significant external financing.

But, despite the difficulties, participants left Paris in optimistic mood. "With the response we've gotten, and Unesco's strong support, I am convinced that we will get the money," says Herman Winick of the Stanford Linear Accelerator Centre in California, one of the project's co-founders. **Heather McCabe**

Israel 'must relax technology transfer law'

[JERUSALEM] Orna Berry, chief scientist in Israel's Ministry of Commerce and Industry, acknowledged last week that the country's law on research and development (R&D) needs to be revised to take into account changes in intellectual property rights and technology exports.

Her statement reflects growing pressure on the government to liberalize its policy on technology transfer, and to lift restrictions requiring the application of government-funded research to take place primarily in Israel.

Berry says that her office is planning to submit to the new Knesset, the Israeli parliament, proposals to modify the R&D law to allow greater flexibility in the exploitation of government-funded research. But many in the country's high-tech industries argue that the proposed changes do not go far enough.

Speaking at a seminar at Ben-Gurion University in Be'ersheva, Berry said that, although she has much flexibility under the current law, she would prefer to see these policies inserted explicitly into the law.

According to government officials, the main problem with intellectual property rights in Israel is that companies are often unaware of the value of their ideas, and sell them overseas too cheaply. One official estimates that there are 30 such cases a year.

Berry argues that the R&D law needs to be changed to defend Israel's intellectual property, but without restricting the marketplace. For example, the law requires that a company receiving an R&D grant from the government must carry out its production in Israel or return the grant. This dissuades some start-up companies that expect to need overseas partnership from applying for government assistance.

Berry supports this requirement as

necessary to job creation in Israel. But she acknowledges that there are cases in which overseas production is more profitable, and argues that the law should be changed to allow for them. It should also help Israeli firms to acquire the managerial and marketing know-how needed to make local production profitable, she said.

Another gap in the law concerns mergers and acquisitions. Companies that receive grants must pay royalties to the government out of their sales. But many companies are bought by foreign companies before there is any production, and the government loses its investment.

Berry wants the law changed so that companies bought out in this way must still repay their grants. Her insistence that most high-tech production should be carried out in Israel was criticized by many industry representatives at the seminar, especially those involved in biotechnology, where investors and industrialists see overseas production and marketing as important elements in developing Israel's potential.

Critics argue that, although Israel's life scientists produce good research, the local management and marketing expertise needed to turn research into commercial products is lacking. "No government policy in the world has made a difference for biotechnology — it all comes from the market," said Ehud Geller of Medica Venture Partners, a biotech venture-capital firm.

Geller said that the Israeli government should let the market do its work, including allowing foreign companies to acquire Israeli technologies.

Geller said pension funds and insurance companies — the largest investors in the country — invest almost nothing in biotech because of government restrictions and negative incentives. **Haim Watzman**

Researcher fights suspension over funds

[SAN DIEGO] The University of California at San Francisco (UCSF) has suspended for three years an internationally known anaesthesiology researcher who diverted millions of dollars from drug company grants to a non-profit foundation that he directed.

But before the academic suspension could begin on 1 June, the researcher, Dennis Mangano, filed a lawsuit in state court in a bid to block the rare academic disciplinary action in the long-running case (see *Nature* 385, 377; 1997).

University attorneys said they had delayed imposing the suspension until after court proceedings because Mangano is on unpaid, stress-related leave from his tenured professorship at UCSF.

But he reportedly continues to practice medicine at the VA Medical Center in San Francisco, where he was a vice-chairman of anaesthesia.

The Mangano case is being closely watched at the University of California, where non-profit foundations set up by academic researchers have come under scrutiny. Such foundations came into vogue about 20 years ago to enable benefactors to contribute to specific research projects, and professors to fund immediate research needs.

A number of foundations created by University of California professors grew to the million-dollar range, and the Ischemia Research and Education Foundation (IREF), founded and directed by Mangano, became one of the biggest.

A UCSF audit estimated that more than \$50 million in indirect costs were diverted to IREF. In his lawsuit, Mangano says IREF collected indirect expenses for possible future needs at an independent research site.

Following academic reviews and a management audit, court records say that UCSF Chancellor Michael Bishop notified Mangano on 3 May of his suspension after a university committee had found that he "repeatedly made extensive use of university resources" to benefit his foundation and for "purposely defrauding sponsors of drug trials". The committee also found that Mangano harassed and intimidated UCSF colleagues.

Mangano's attorneys declined to be interviewed on the suspension or the lawsuit, which was filed on 28 May in San Francisco Superior Court. They have sought to keep sealed the 65-page lawsuit detailing the five years of the academic disciplinary process and fighting over millions of dollars in research funds.

But *Nature* obtained a copy of the lawsuit after it was revealed at this month's meeting of the University of California Board of Regents.

Objecting to the suspension, Mangano's attorneys say in court records that he was

denied due process, that the disciplinary academic committee's decisions are not supported by the evidence, and that university rules regarding non-profit foundations were not violated. The lawsuit seeks to have the suspension voided and Mangano paid damages for what his attorneys called "this virtually unprecedented discipline".

"The suspension will devastate [his] academic career and will damage his reputation in the professional community beyond repair," the lawsuit says. The university has until next week to respond.

In January 1997, IREF paid \$25 million to settle a lawsuit over the alleged misappropriation of indirect costs associated with international drug studies conducted by Mangano and IREF.

The lawsuit had been filed in 1994 by physiologist Charles A. Richardson, a former colleague of Mangano's, invoking a seldom-used California law for whistleblowers to expose misappropriations from the state.

The university received \$19.2 million of the settlement, in which IREF admitted no wrongdoing. Richardson got \$1.5 million and his attorneys \$4.3 million; payments the whistleblower law permits to encourage the reporting of potential improprieties.

But in a peculiar twist, Richardson lost his job as a UCSF assistant adjunct professor less than a year later. Last February, Richardson filed a lawsuit against the university and UCSF's chairman of anaesthesia, Ronald D. Miller, for retaliation, alleging he was forced out for challenging practices associated with

non-profit foundations.

Richardson sought to have some of the money recovered from IREF used to support his research projects, which were disrupted when he challenged Mangano. University officials, who rejected Richardson's request, have denied any retaliation.

During Richardson's lawsuit and his most recent case, Mangano argued that that his non-profit corporation activities were no different to other UCSF professors. Apparently to support this, Mangano's attorney persuaded Miller to acknowledge during a deposition that the anaesthesia chief used \$17,000 from a foundation with which he was associated to renovate his UCSF office. In his deposition, Miller said he intended to pay the foundation back. Miller could not be reached for comment last week.

John Lundberg, the university's deputy general counsel, says that the suspension is appropriate because Mangano's "conduct was beyond the pale".

Lundberg said a tactical decision was made to settle the whistleblower's lawsuit in 1997 rather than fight for the remainder of the money. Most of the other foundations started by University of California professors have been required to direct their funds to the university's foundations.

Since the Richardson lawsuit, Mangano has been involved in several other lawsuits over the handling of IREF funds. In one, he is accused of improperly using more than \$300,000, which he denies. He now claims the foundation owes him \$1.4 million. **Rex Dalton**

First-born telescopic twin starts to deliver



[LONDON] Astronomers have begun to receive pictures from one of the twin infrared telescopes, Gemini North (above), based at Mauna Kea in Hawaii, which is being dedicated this week (25 June). One of the first images is of the planet Pluto and its moon Charon (inset).

The ground-based infrared telescope is more than twice as powerful as the Hubble space telescope and has an eight-metre mirror. Yet the twin Gemini telescopes will

cost only £111m (US\$176 million), a fraction of Hubble's \$3 billion.

The project is an international effort, with most support coming from the United States and Britain. Gemini South is under construction at Cerro Pachon, Chile.

Gemini is able to partially correct for atmospheric distortions producing high-quality images. On Monday this week, astronomers received a picture with a resolution of 0.08 arcseconds. **Natasha Loder**

Japanese body calls for ban on research into human cloning

[TOKYO] Japan has moved closer to banning research involving human cloning, after the Council for Science and Technology — the country's main science policy-making body — last week proposed that such research should be regulated by law.

The council's cloning subcommittee said that the most appropriate way of regulating human cloning research would be to impose a legal ban, and that research using cloned human embryos less than 14 days old should be controlled by means of guidelines, given its potential medical applications.

French charity scandal trial draws to a close

[PARIS] The French public prosecutor last week called for five-year prison terms for Jacques Crozemarie, founder and ex-president of the French Association for Cancer Research, and Michel Simon, former chief executive officer of International Development, a public relations company.

The two men are charged with embezzling up to FF300 million (US\$47 million) of the charity's funds for personal gain (see

Nature 379, 103; 1996). Speaking at the trial, which is due to end next week, Paulette Arrault, the representative of the public prosecutor, accused Crozemarie of "organizing and premeditating the pillage" of the charity "to assure his personal glory, and satisfy his need to be seen".

NIH targets the right diseases, study shows

[WASHINGTON] Spending by the US National Institutes of Health (NIH) on specific diseases is "associated with" the burden of the disease to society, according to a study in last week's *New England Journal of Medicine*. A measure called the disability-adjusted-life-year — defined as the loss of one year of healthy life to disease — was found to be "strongly predictive" of NIH spending on 29 diseases in 1996.

NIH director Harold Varmus wrote in an accompanying editorial that the study confirms the NIH's claim that its spending shows a "reasonable correspondence" with the burden of disease. But he added: "It is important to emphasize that there is not, and should not be, an absolute correspondence."

Varmus said that other factors, such as scientific opportunity and the amount of industry research on certain diseases, also affect the agency's spending decisions.

US genome institute triples sequencing rate

[SAN DIEGO] The US Department of Energy's Joint Genome Institute in Walnut Creek, California, has tripled its capacity to decode human DNA by leasing 24 MegaBACE DNA sequencers, officials announced last week. Operating at a newly dedicated facility, the Joint Genome Institute — which is targeting human chromosomes 5, 16 and 19 — now has the capacity to sequence 14 million raw bases per day.

Federal officials say they plan to acquire more sequencers to increase the decoding effort further by a factor of two or more. More than 200 researchers now work around the clock sequencing DNA at the institute, with their efforts published daily on a website.

Bulgaria's particle physicists join CERN

[GENEVA] Bulgaria has become the twentieth member state of the European Laboratory for Particle Physics (CERN). At a ceremony in Geneva last week, Vesselin Metodiev, Bulgaria's deputy prime minister and minister for education and science, said that one of his government's priorities "is to develop and maintain competitive science in Bulgaria" and that "membership of CERN is

an important step in this direction”.

Bulgaria has small theoretical and experimental particle-physics communities, mainly at the Institute of Nuclear Research and Nuclear Energy of the Bulgarian Academy of Sciences and at the University of Sofia. Some research groups already participate in CERN experiments.

German fraud case finally comes to court

[MUNICH] More than two years after the emergence of Europe's biggest case of alleged scientific fraud, two German scientists — Friedhelm Herrmann, a clinical researcher, and Marion Brach, a molecular biologist — are finally being brought to court in Ulm, where they had worked at the local university.

Herrmann and Brach are charged with using publications based on manipulated data to support their applications for posts as university professors (see *Nature* 387, 442; 1997). “This constitutes fraud,” argues public prosecutor Konrad Menz.

The *Land* (state) of Baden-Württemberg, their formal employer, is seeking damages equivalent to the salaries paid to Herrmann and Brach during their employment at the University of Ulm. If found guilty, Herrmann and Brach could receive prison sentences of “between five days and five years”, says Menz.

Website launched for women in science

[LONDON] A web-based resource centre for women in science was launched at Britain's Sheffield Hallam University this week. The centre has been set up by the group Women in Science, Engineering and Technology (WiTEC) as a one-stop shop of information relevant to women's career development in science, engineering and technology.

The database links to similar databases compiled by WiTEC's European counterparts, and has two years of support from the UK Department of Trade and Industry and the European Commission's Leonardo Da Vinci fund. The web address is: <http://ntcrater.adc.shu.ac.uk/>

Customs stop mammoth smuggling operation

[MOSCOW] A truck containing around 1,000 kg of palaeontological specimens was detained by customs officials in Smolensk last week. The haul, estimated by Russian police to be worth more than US\$1 million, was found to contain complete skeletons and skulls of baby mammoths and cave bears, and mammoth tusks.

Described on export documentation as “bone fragments of no scientific or

commercial value”, the load was being driven over the Latvian border to a destination in Germany. This is the second time in six months that police have foiled an attempt to smuggle a large haul of scientifically important fossils, which cannot legally be exported (see *Nature* 397, 189; 1999).

Miledi and Moreno win Spanish science award

[BARCELONA] Neurobiologist Ricardo Miledi and surgeon Enrique Moreno have been awarded the Príncipe de Asturias, Spain's top award for scientific and technological research, worth 5 million pesetas (US\$33,000). The award panel wanted to draw attention to the collaborative nature of biomedical research today, with its basic, clinical and technological sides, said its president Julio Rodríguez-Villanueva.

The panel said that Miledi's studies and discoveries in neurophysiology, especially in the molecular characterization of neurotransmission receptors, were pivotal in understanding the information transmission pathways in the nervous system. Miledi was born in Mexico. Moreno is the author of important contributions in abdominal oncology and transplant surgery. He has been heavily involved in helping to increase organ donation rates in Spain.

nature

the international weekly journal of science

needs a Web Editor

Are you seriously fascinated by the Web?

Do you seriously love science?

Would you like to be in charge of the Web presentation of the world's most innovative and, sometimes, publicly controversial science?

Do you have strong editorial experience and flair? Do you have a solid scientific background at least to graduate level?

If your answer is “yes” to all of these, then you could be the person we are looking for. *Nature*, the world's leading international weekly journal of science, needs a Web Editor to take the lead in our development of online services and content.

Our site is already heavily visited by scientists worldwide. The challenges would include the development and launch of new services from the group of Nature titles online: *Nature* itself, and six related monthly publications. You would also lead editorial innovation in presenting our content, week by week and archivally, on the Web. And we'd want strong input to strategic thinking about online science publishing across all



disciplines, in a highly competitive and fast-changing environment.

It's a serious job — but stimulating, too.

After all, we want www.nature.com to be THE place for the world's liveliest and most demanding scientists.

This is a senior post, based in our London office, reporting to the Editor of *Nature*. Salary and terms are highly competitive.

Applicants should send a CV, examples of relevant work, and a brief summary (1000-word maximum) of strengths and weaknesses of Nature's web site, and their priorities for its development. Mark the envelope “Web Editor”, and send to **Dr Philip Campbell, The Editor, *Nature*, 4 Crinan St, London N1 9XW**, to arrive no later than Monday 12 July. Applications may also be sent by e-mail to webeditor@nature.com.



The full text of items on this page – and further details and background information about the Unesco/ICSU World Conference on Science, to be held in Budapest from 26 June to 1 July – can be found on *Nature's* website at <http://www.nature.com>

Build confidence in research advice, says UK chief scientist

[LONDON] Suggestions from the World Conference on Science on developing policies for science should be complemented by attention to the parallel task of building science into wider government policies, according to Sir Robert May, Britain's chief scientific adviser.

According to May, the public must have confidence in the way that science advice feeds into policy making. Policies must therefore be developed to take account of the benefits and opportunities that stem from advances in scientific knowledge, as well as caution about the potential risks.

Scientific specialists should join with others in open and wide-ranging consultations that identify, rather than exclude, particular interests and conflicts of interest, and place them in a broader perspective.

"At the same time, the fact that science, and the major issues that science and technology address, are becoming increasingly international, means that such issues need to be tackled co-operatively by all nations," says May.

May urges governments to develop their procedures for incorporating scientific advice into policy making. They should establish ways of enabling the early anticipation and identification of issues for which scientific advice or research will be needed.

Governments should be able to draw on a range of expert sources, and involve them in framing and assessing policy options. Experts from other countries or international advisory mechanisms should be involved wherever possible.

"The challenge for the Budapest conference is to focus on reaching practical proposals on these and other issues," says May. "It must not set over-ambitious objectives, while whatever emerges should be of practical value, identifying the extent and cost implications of any commitments reached."

He warns that implementing elements of the proposed draft *Framework for Action* could lead to the creation of "unnecessarily duplicative" follow-up groups. Given the diverse aspirations and resources of the participants, however, it is unlikely that any participant will attempt everything that is set out in the *Framework*, says May.

"Rather, they are likely to adopt a piecemeal approach, a bit like selecting from a menu. The task of those participating in the Budapest meeting is to ensure that this menu is both tasty and nutritious."

Full text: <http://helix.nature.com/wcs/c22.html>

Readers give their hopes and fears for conference

[LONDON] Science ethics, environmental issues, poverty, privatization of knowledge, and bridging the knowledge gap between countries, are the priorities that readers of *Nature's* website would like to see addressed at next week's World Conference on Science.

But many are worried that the conference, although studded with good intentions, may not lead to practical action. The single most common concern is that the conference will be "all talk, but no action". Very few readers believe that it will result in more funds for developing countries.

Almost 300 readers of *Nature's* World Conference on Science web pages have already responded to a survey in which they were asked about their priorities and concerns for next week's conference in Budapest.

Asked to name the top priorities that the conference should address, a large majority of respondents listed at least one of the following: free access to knowledge; biotechnology ethics; using science to address poverty, health and environmental concerns; science education; and public access to science.

Many are concerned about the lack of adequate controls on science. But many — sometimes the same respondents — hold out high hopes that science can play a leading role in reducing poverty and in tackling environmental issues such as ozone depletion, loss of forests, and climate change.

One reader from the Netherlands said he wants the conference to address: "The role of science in defining and achieving sustainable world development; and how to strike a balance between scientific freedom and the demands for certain ethical controls on what science is done and how."

Another reader, from India, said he wants

the conference to address: "How science should come to terms with other knowledge systems and other civilizational systems."

Many respondents called on scientists to take more responsibility for their work and to become better public communicators. Those respondents who are scientists included in their priorities: difficulties in securing permanent jobs; insufficient funds for basic research; the need for more women scientists; and better facilities and access to science in developing countries.

Asked to name their main concern about the conference, an overwhelming majority of respondents said they feared it would be little more than a talking shop. "A wishy-washy, slap-on-the-back, aren't we great affair," is how one reader fears the conference might turn out. "One's main worry is that it is a lot of hot air leading nowhere," remarked another reader, from Australia.

Some believe the conference is dominated by government-appointed delegates, and should have included more representatives of young people and non-governmental organizations, from all countries.

"I would like to participate, even paying myself, but was told that a special invitation is required. Who can give me one?" asks one reader from Italy. A reader from the United States expresses concern that "It will be simply a gathering of the 'usual suspects', speaking politely and guardedly, rather than facing crucial issues squarely and speaking frankly."

But one reader from Belgium is a little more optimistic: "I think it is a very good starting point to think about science and its impact on society. The risk is that there will be no concrete actions, only nice intentions."

Full text: <http://helix.nature.com/wcs/a47.html>

Nature brings daily news from Budapest

Nature's news writers will be providing daily electronic coverage of the World Conference on Science from Saturday 26 June, the first day of the meeting, up to Friday 2 July, the day after the closing session.

Material will include news items on the main speeches delivered to the conference, reports on discussions about the texts of the *Declaration and Framework for Action* that are due to be adopted on the final day of the conference, and interviews with key participants.

A questionnaire that includes the

opportunity for readers to submit their hopes and fears for the conference (see above) remains open on the website, which can be reached either through the *Nature* home page (<http://www.nature.com>) or direct on <http://helix.nature.com/wcs>.

All respondents to the questionnaire will be eligible to receive a free supplement that *Nature* will be producing after the conference. This will include reports on the meeting and its conclusions, as well as selected opinion articles on the topics under discussion in Budapest that have appeared on the website.

Devolution threat to decision-making

Sir — The issues underlying the symbiosis of the institutions of civil and scientific administration are crucial for the future of society, and are addressed in a masterly report on risk by the Royal Society¹, which has probably been read by less than 0.1 per cent of the UK population. In this ill-defined context, the new parliaments in Scotland and Wales have begun a process of devolving science policy and its administration.

Nature, the Royal Society and the pressure group Save British Science (SBS) have all considered the problem^{2,3}. What is emerging seems to be a determination to equip the new parliaments with new ministers for science, new supporting bureaucracies, and what has been described optimistically as a "science champion". An "extra" £300 million (US\$477 million) has been promised by the UK government to Scottish science over three years. The SBS implies that Wales is more conscious than Scotland of the fact that the United Kingdom can hardly be described as an unmanageably large base for scientific endeavour and its administration. This indicates the usual obeisance that political parties pay to pressure groups.

Nothing, however, could be more disastrous for UK science than a fragmentation of the decision-making

system. The country certainly does not need a separate Welsh or Scottish Parliamentary Office of Science and Technology (POST). It does need to ensure that the services of POST are made available to both new parliaments. Any additional funding required should be contributed by the new parliaments.

New arrangements will not increase the resources of the institutions determining science policy in the United Kingdom as a whole. The £300 million allotted to Scotland will inevitably reduce the science budget elsewhere in the United Kingdom, unless the overall budget is increased.

Europe already has several 'national' POSTs, most of which were modelled on the UK POST or the US Office of Technology Assessment. The European Parliament has established its own assessment mechanism (STOA). All depend on a limited range of independent scientific advice which should not be influenced by the nationality or location of the scientists involved. If policy judgements are now to depend on scientific advice to governments, or to legislatures required to endorse their decisions, the only relevant criterion is the integrity of the analyses and those who contribute to them. There is never an abundance of independent expertise and the more novel the subject of the enquiry,

the greater the difficulty of establishing a 'peer group' whose partiality is not seen to be suspect by interested parties

The controversy over genetically modified foods exemplifies the need for public confidence in scientific opinion. Disagreements should be resolved within the scientific community and not be presented as Scottish, Welsh, UK or any other 'national' opinion. Good science needs no national or sub-national prefix.

The UK parliament, for all its faults, was the first legislature to establish (in 1938) a Parliamentary and Scientific Committee. Its science and technology select committees have ranged over innumerable controversial issues. The political problems generated by the exponential advance of scientific knowledge are those that transcend national boundaries. They will not be solved by a fragmentation of the scientific advice provided to the community in an ill-conceived expansion designed to appease national egos inflamed by political ambition. The fundamental internationalism of science must not be damaged, within or outside the United Kingdom.

Sir Ian Lloyd

Bakers House, Priors Dean, Petersfield GU32 1BS, UK

1. *Risk: Analysis, Perception and Management* (Royal Society, London, 1993).

2. *Nature* **399**, 89 (1999).

3. Masood, E. *Nature* **399**, 97 (1999).

Farm-scale evaluation of GM crops explained

Sir — The first genetically modified (GM) crops being proposed for commercial planting in the United Kingdom have been altered to make them less sensitive to broad-spectrum herbicides. These crops are intended to allow more efficient weed management and herbicide regimes for the farmer, reduced frequency and quantity of herbicide applications, and increased market share for the suppliers. Work has now begun on farm-scale studies of maize and spring oil-seed rape, with winter oil-seed rape studies starting later this year. Many objections to GM crops have been raised, and there is pressure on farmers not to take part in this research programme (see *Nature* **398**, 651–656; 1999). Indeed, one farm, near Swindon in the west of England, has recently withdrawn from the study.

Because of the widespread public interest in GM crops, we, as representatives of the consortium commissioned to conduct this research, think it is worthwhile to outline the background and purpose of

these farm-scale evaluations. The research addresses the concern that the changing management of the GM herbicide-tolerant (GMHT) crops could result in reductions of weed and invertebrate populations on which farmland birds and other wildlife depend. This concern has been voiced most cogently by the conservation groups English Nature and the Royal Society for the Protection of Birds (RSPB).

We are aiming to test the null hypothesis that there are no differences between the biodiversity associated with the growing of GMHT crops and comparable non-GM crops at the farm scale. The study will look for positive and negative effects. These are likely to be indirect, resulting from crop and rotation management, rather than from a direct effect of the use of GM plant breeding technology. Indeed, if herbicide resistance had been introduced by traditional breeding, the design of the study would have been the same. Farmers will grow and manage both GM and non-GM crops as they would do commercially.

The only constraints are that the varieties being compared are as similar as possible in other characteristics, and also that all other agronomic practices are kept

the same, unless commercial reasons dictate otherwise. So we anticipate that the herbicide regimes will differ, but that any insecticide or fungicide should be applied to both treatments on each farm on the same day, unless there are clear agronomic reasons that may in themselves be the result of growing the different crops. We will also evaluate effects on biodiversity in following crops. These crops will be the same, and will be managed in the same way, following both GM and non-GM treatments, unless there are clear agronomic reasons for differences in rotations or management.

It is not practical to record population responses for all species in the arable system. We are using indicators of biodiversity that are likely to be sensitive to the treatments, and reflect processes that may lead to significant ecological shifts that cannot be detected directly given the time and spatial scales available for the study. Effort will be concentrated on species groups that do not forage over wide areas or occupy higher trophic levels. These include vegetation in and around the crop, the field seed bank, earthworms, snails and slugs, ground beetles, bugs (Heteroptera), foliage arthropods and invertebrate biomass (with

moth and sawfly larvae measured separately). Bees and butterflies will also be assessed. These are all being recorded using standard protocols that are being tested and refined during this year.

Birds are not included in the field study because they range too widely to show real effects when only single fields are being considered, although data on invertebrates and plants will provide measures of resources available to them.

Work in the first season is on a pilot scale to ensure that monitoring is matched to the details of crop management. There will then be three seasons of the summer crops and at least two of the winter crop, at the full scale of around 20 treatment pairs per crop per year. We will select from the pool of available farms using a stratified random procedure; the experimental treatments will be allocated at random within each farm. The GM and control crops will be grown in a split-field or a paired-field plan; work in this first year will confirm which is the more appropriate. There are valid arguments for and against both configurations. In a split field, the two halves of the field will have had similar histories, reducing the variation in biodiversity indicators before treatment. The paired-field design gives less chance of interference between the treatments and is more realistic in terms of the structure of the field boundaries. Both configurations are included in the first-year sites.

The work is being conducted by a UK consortium of the Institute of Terrestrial Ecology, the Institute of Arable Crops Research and the Scottish Crop Research Institute. It is funded by the Department of the Environment, Transport and the Regions, the Ministry of Agriculture, Fisheries and Food, and the Scottish Office.

A steering committee will oversee the progress of the work and ensure the scientific quality and integrity of results. The committee includes independent scientists, including experts from English Nature, RSPB and the Game Conservancy Trust. Many results will not be available until the end of the project in 2002.

The role of scientists is to provide the evidence on which to base a sound risk assessment of the effects of herbicide-tolerant GM crops on biodiversity.

Our evidence will, we trust, provide an important input into a rational debate about the adoption of GM crops.

L. G. Firbank*, **A. M. Dewar†**, **M. O. Hill‡**,
M. J. May†, **J. N. Perry§**, **P. Rothery‡**,
G. R. Squire¶, **I. P. Woiwod§**

*Institute of Terrestrial Ecology, Merlewood Research Station, Grange-over-Sands LA11 6JU, UK

†Institute of Arable Crops Research, Broom's Barn

‡Institute of Terrestrial Ecology, Monks Wood

§Institute of Arable Crops Research, Rothamsted

¶Scottish Crop Research Institute

Bioethicists must come down to Earth

Sir— You report, without critique, the opinion of Canadian bioethicist Margaret Somerville that “science will need to wait and to help ethics to catch up” (*Nature* **399**, 12; 1999). Any regular reader of your journal is sure to wonder on what planet Somerville has grown up. It is certainly not one on which science or private industry exist, for if it was she would surely know the lunacy of her proposition.

We would be better served if bioethicists were willing and able to work within the realm of the modern, market-oriented world to come up with practical solutions to bioethical problems.

Lee M. Silver

Department of Molecular Biology, and Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, New Jersey 08544, USA

Barking up the wrong pole

Sir— In a review of Freeman Dyson's book, *The Sun, the Genome and the Internet, Tools of Scientific Revolutions*, the reviewer writes of “such extravagances as bringing back lumps of rock from Mars, when nature has already left us generous supplies of the same material in the form of meteorites, mostly still reposing in the Arctic ice” (*Nature* **398**, 770; 1999). I assume he is in fact referring to the Antarctic blue ice meteorite recovery areas such as Lewis Cliff, Antarctica.

It could be possible to recover more than 100,000 meteorites in the Antarctic over the next couple of decades. In 1986–87, for example, we recovered several hundred meteorites.

Austin Mardon

Antarctic Institute of Canada, PO Box 1223, Main Post Office, Edmonton, Alberta, Canada T5J 2M4

Let's all speak the same language

Sir— In your article on the fifth conference of the African Academy of Science, Ali Mazrui is reported as suggesting that African science is unlikely to develop while English remains the main medium of communication (*Nature* **399**, 12; 1999). I can understand the desire to discuss one's work in one's own language, but I must

question how practical it would be in a continent such as Africa where there are many indigenous languages.

There have been several successful pan-African conferences on natural products chemistry, a subject which I think Mazrui would consider valuable, in view of the scope that it offers for examining traditional medical knowledge. I could not help picturing what such a conference would be like if conducted in African languages, with simultaneous translation into Arabic, Amharic, Swahili, Yoruba and Zulu.

Even deciding which to accept as official conference languages might provoke endless disagreement. Possibly, Mazrui's suggestion might be more appropriate in the romantic field of the literary world, rather than in the more practical scientific one. However nice as an idea, I think that the suggestion has little practical relevance.

D. A. H. Taylor

12 Avenue Road, Scarborough YO12 5JX, UK

German researchers won't be put in the dock

Sir— Your article “Animal rights activists turn the screw” stated that the Deutsche Tierschutzbund [a German animal welfare organization] would “initiate court cases and injunctions against researchers” if animal protection were included in the German constitution (*Nature* **396**, 505; 1998). Contrary to this statement, the Deutsche Tierschutzbund has no intention of doing so.

The proposed change to the constitution aims to reinforce a 1986 amendment to the German animal welfare law that introduced a requirement for licensing procedures for experiments to include an ethical evaluation process. The need for this change arose after the Constitutional Court decided in 1994 that such ethical evaluation is unconstitutional, because freedom of research is embodied in the constitution, but animal welfare is not.

No animal welfare organization had brought a court case against researchers before 1994, so why should this change if the requirement for ethical evaluation is simply reinforced? Animal welfare organizations will find it hard to take scientists to court or to have licences revoked: the licensing procedures will remain confidential, and the decision of the authorities will rest on criteria that are not heard at court.

Wolfgang Apel

(President)

Deutsche Tierschutzbund,

Baumschulallee 15, 53115 Bonn, Germany

The end of the road for silicon?

Max Schulz

Computer chips continue to shrink. But the discovery that a layer of silicon dioxide must be at least four to five atoms thick to function as an insulator suggests that silicon-based microchips will reach the physical limits of miniaturization early next century.

Silicon-based microelectronic devices have revolutionized our world in the past three decades. Integrated circuits, built up from many silicon devices (such as transistors and diodes) on a single chip, control everything from cars to telephones, not to mention the Internet. The thirst for cheaper electronic memory, and faster and more powerful processors, is still not satisfied. Each year we see more powerful chips with smaller device features, making them smarter and cheaper. The miniaturization of the devices found in integrated circuits is predicted by the semiconductor industry roadmap to reach atomic dimensions in 2012. According to Muller *et al.* (page 758 of this issue¹), the narrowest feature of silicon devices—the gate oxide—will then reach its fundamental physical limit. In a transistor, the gate oxide insulates the voltage electrode from the current-carrying electrodes (Fig. 1). At a thickness of less than four layers of silicon atoms, current will penetrate through the gate oxide causing the chip to fail.

In 1925, Lilienfeld patented² the first field-effect device (one where current flow is modified by applying an electric field) based on silicon, but he probably never got it to work. It wasn't until 1960 that Kahng and Atalla³ demonstrated the first metal-oxide semiconductor field-effect transistor (MOSFET), a remarkably simple device

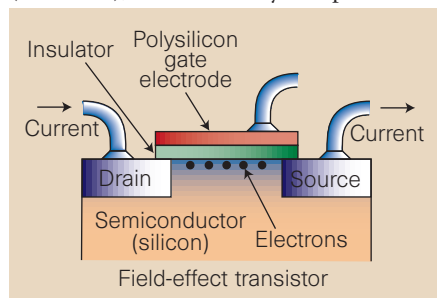


Figure 1 A field effect transistor (FET), such as that used in computer memory chips. The FET consists of source and drain channel contacts, and a polysilicon gate electrode separated from the semiconductor silicon by the insulator SiO_2 . When the voltage on the gate is positive, electrons accumulate on the semiconductor surface making the channel between source and drain conducting and so turning the transistor from the 'off' into the 'on' state.

(Fig. 1), in which the semiconductor silicon plays a crucial role. This planar electronic device revolutionized electronics because a large number of MOSFETs and their interconnections could be built up on the surface of a single silicon chip.

Only ten years later, the first integrated circuits—a 1 kilobit memory chip and a 750 kHz microprocessor unit—appeared on the market as the first 'large-scale integrated' devices with 100 to 5,000 components squeezed onto a single chip. Since then, continuous evolution has increased the number of device components on a chip by a factor of 64,000 to a fully integrated 64 megabit memory chip, in which there are more than one hundred million electronic components. The scaling down of device sizes not only increases the number of transistors per chip, but also increases the speed of the circuits, up to 600 MHz in today's personal computers.

The progress up to now is well described by 'Moore's law'⁴. Gordon Moore predicted

in 1965 that for each new generation of memory chip and microprocessor unit on the market, the device size would reduce by 33%, the chip size would increase by 50%, and the number of components on a chip would quadruple every three years. So far this trend has shown no sign of stopping.

Several properties of silicon have made these developments in microelectronics possible. Silicon can be grown in single crystals more than 1 m long and 30 cm across, weighing approximately 200 kg. The purity of the crystal and the number of electrically active defects are well under control. The number of atomic crystal defects in sub-micrometre-sized MOSFETs is now limited to individual centres that act as traps for electrons. Such traps may be identified, individually characterized, and counted, so that single-electron transistors are possible⁵.

The special feature of silicon, which makes it the semiconductor of choice for MOSFETs, is its native oxide. Silicon dioxide (SiO_2) is an almost perfect insulator with a resistivity in excess of $10^{16} \Omega\text{cm}$. The insulating films of SiO_2 grown on silicon are smooth and coherent with no holes in a thickness range down to single atomic layers^{6–8}. The interface with silicon is abrupt and there are very few electrically active defects at the interface⁵. In the laboratory it is now possible to produce MOSFETs and integrated circuits⁹ with gate oxides less than ten atoms across. Such thin films are required to maintain the current response of the transistor to lower voltages at the gate electrode. Manufacturers need to lower the

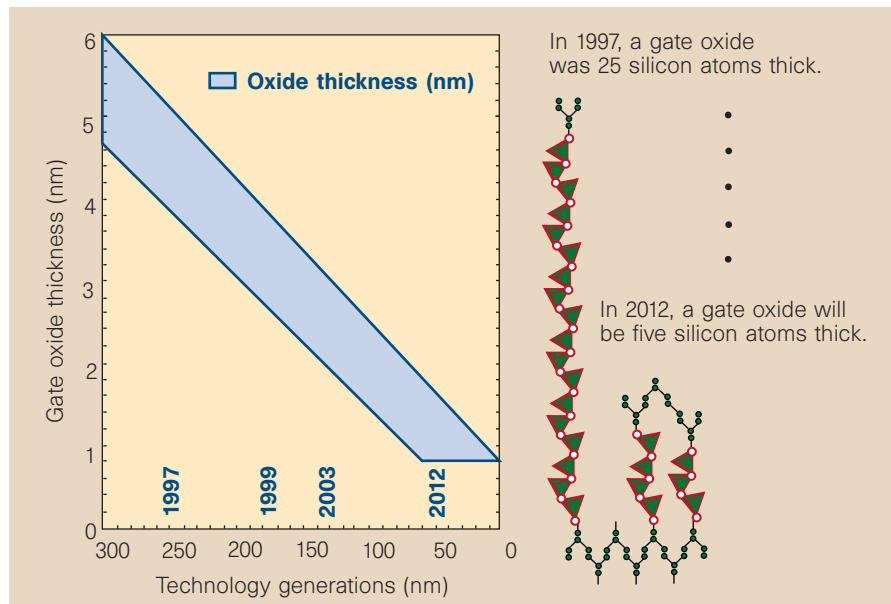


Figure 2 Semiconductor industry roadmap. Predictions of the gate oxide (SiO_2) thickness for future technology generations, which are defined by the critical device size. The gate oxide is so thin that it can be drawn on an atomic scale (see right of plot). Full circles indicate the silicon atoms in the silicon substrate and the polysilicon gate electrode. White circles indicate the silicon atoms in the oxide structure. In 2012 a gate oxide will be only 1.3 nm thick, or five silicon layers thick. Two of the silicon atoms are bound to the crystalline silicon at the interface, so the 'bulk' oxide only consists of three atomic layers, which is now demonstrated by Muller *et al.*¹ to be just enough to provide a working insulating layer.

power supply to individual components, if they are going to cram more devices onto a chip. The performance of the gate oxide is crucial to the device, so ultrathin gate oxides are being studied intensively, as witnessed last week at a conference on insulating films on semiconductors in Kloster Banz, Germany¹⁰.

The Semiconductor Industry Association in the United States extrapolates Moore's law into the future to create a roadmap that predicts future developments in microelectronics based on silicon devices¹¹. The roadmap sets yearly targets for the performance of integrated circuits and defines the advances in semiconductor technology required to reach those goals. Figure 2 shows the 1997 prediction for the next few generations of chips. In this plot the generations are defined by a critical device size, which is projected to decrease from 200 nm to 50 nm over the next 12 years. The gate oxide must be reduced in turn, from 25 silicon atoms today to 5 atoms in 2012 to achieve the roadmap goal. Clearly, there must be a limit to this scaling down because the gate-oxide thickness will eventually reach zero. The end of the line for the SiO₂ gate insulator is now expected early next century.

The work of Muller *et al.*¹ from Lucent Technologies is an experimental proof of the fundamental physical limit to the size of a working gate oxide. Using an electron microscope in combination with spectroscopic analysis of the electron energy levels in the Si-SiO₂ interface, they show that electronic wavefunctions penetrate through the ultrathin oxide film from both interfaces. This means that the electrical insulation of the gate oxide breaks down for an oxide thickness less than 0.7 nm, or four atomic layers of silicon in the oxide. Such an oxide is actually only two atomic layers thick, because the two silicon atoms at the boundaries of the oxide are not completely oxidized (Fig. 2). In practice the gate oxide has to be slightly thicker than the theoretical limit, say five layers of silicon atoms, because the interface is rough on an atomic level.

In the past, many technological limitations to Moore's law have been predicted. All these have been overcome by new developments, leading to silicon materials of unprecedented quality. Device processing is carried out in clean rooms with less than one dust particle per cubic metre. The components of an integrated circuit are usually fabricated by lithography — etching a pattern onto the silicon using blue light. As device size continues to shrink, lithographic techniques have moved to shorter wavelengths of light, with chips today being manufactured using ultraviolet light sources (193 nm). Future devices will require structures smaller than the wavelength of the light used, but there may be a way round this problem by using phase-contrast methods. So far there has seemed to be no obstacle to stop the

continuous development of microelectronic chips.

The new limit to oxide thickness is fundamental, however, and cannot simply be overcome by technological improvements. The end of SiO₂ as a gate insulator will be reached in the year 2012, according to the roadmap for silicon technology. The science community and the semiconductor industry will have to come up with new ideas to avoid a bottleneck in growth. One possible solution is to use an insulating film with a dielectric constant higher than that for SiO₂. A high dielectric constant allows the use of a thicker insulating layer for the gate electrode. Unfortunately, all the materials studied so far are not as good as SiO₂. It is going to be difficult to replace SiO₂ or to modify its composition to increase the dielectric constant. But, so far, all the problems in the development of semiconductor technology have been solved. I am almost certain that this barrier will also be overcome, but this time using a different solution than further thinning of the oxide. Perhaps new device structures more complex than the MOSFET will one day be needed¹².

For the next 12 years, microelectronic circuits based on silicon technology will remain the basis for further 'electronization' of our

world. The everyday demand for electronic equipment can be met in the short term. Silicon technology is still the most reliable and cost-efficient way to fabricate large microelectronic circuits. It is almost unthinkable that we are nearing the end of the silicon era, but in the meantime silicon chips will be further enhanced. Eventually science and industry will have to find new ways to build faster and larger computers. □

Max Schulz is at the Institute of Applied Physics, University of Erlangen-Nürnberg, Staudtstraße 7, D-91058 Erlangen, Germany.

e-mail: Max.Schulz@physik.uni-erlangen.de

1. Muller, D. A. *et al.* *Nature* **399**, 758–761 (1999).
2. Lilienfeld, J. E. US patent 174175 (1925).
3. Kahng, D. & Atalla, M. US patents 3206670 & 3102230 (1960).
4. Moore, G. E. in *Int. Elec. Devices Mtg (IEDM) Technical Digest* 75 11 (IEEE, Piscataway, New Jersey, 1975).
5. Mueller, H. H. & Schulz, M. *J. Appl. Phys.* **83**, 1734–1741 (1998).
6. Nohira, H., Omura, A., Katayama, M. & Hattori, T. *Appl. Surf. Sci.* **123/124**, 546–549 (1998).
7. Gurevich, A. B. *et al.* *Appl. Phys. Lett.* **74**, 1257–1259 (1999).
8. Timp, G. in *Int. Elec. Devices Mtg (IEDM) Technical Digest* 98 1041 (IEEE, Piscataway, New Jersey, 1998).
9. Momose, H. S. *et al.* in *Int. Elec. Devices Mtg (IEDM) Technical Digest* 94 593 (IEEE, Piscataway, New Jersey, 1994).
10. <http://www.physik.uni-erlangen.de/lap/INFOS99/index.htm>
11. Semiconductor Industry Association *National Technology Roadmap for Semiconductors* 1997 (Sematech Inc., Austin, Texas 1997); <http://www.sematech.org>
12. Nakazato, K., Itoh, K., Mizuta, H. & Ahmed, H. *Electron. Lett.* **35**, 848–850 (1999).

Hox clusters

Size doesn't matter

Mark Q. Martindale and Matthew J. Kourakis

Molecular biology has had a huge effect on our ability to understand the evolution of biological diversity. As we learn more about how embryos develop at the molecular level, we hope to be able to reconstruct the transitions in developmental programmes that lead to new forms. We need to know, for example, which organisms are most closely related to which others; what features these organisms have inherited from a common ancestor; at which critical junctures new innovations appeared; and, finally, how the genetic machinery could allow such transitions. But before we can determine the direction of evolutionary change, we need to sort out the natural evolutionary relationships of living animals. This is the subject of a report by de Rosa *et al.*¹ on page 772 of this issue.

Our understanding of the relationships between metazoan (multicellular) animals comes mainly from detailed examinations of how different animals are put together — the kinds of tissues and organs they possess, and the extent to which their embryonic development is shared. Animals that have a common anatomical organization, or 'body plan', are grouped into 'phyla'. Crustaceans and insects, for example, are invertebrates grouped in the same phylum (Arthropoda).

This is in part because all have individual body segments covered in a hard exoskeleton. Phyla are often then grouped into larger categories, or taxonomic units, such as the 'Articulata' (invertebrate animals with sequential body segments). But these taxonomic units are intellectual constructs, which may or may not reflect features of common ancestry. Even defining concepts such as 'body plan' are difficult to agree on. Instead, an accurate representation of evolution must identify true clades of animals (groups derived from a common evolutionary ancestor).

Molecular biology promises to be a key tool in unravelling such relationships. Using this technique, de Rosa and colleagues¹ have sought to confirm a radical reorganization in metazoan evolutionary relationships. Two years ago, Aguinaldo *et al.*² compared the sequences of a common structural RNA from many different animals with a uniform pattern of molecular evolution (that is, species with unusually fast rates of evolution were eliminated). These authors confirmed the idea that bilaterally symmetrical animals (the Bilateria) can be divided into two groups, deuterostomes and protostomes, which were originally defined by the embryological origin of the mouth (Fig. 1). The

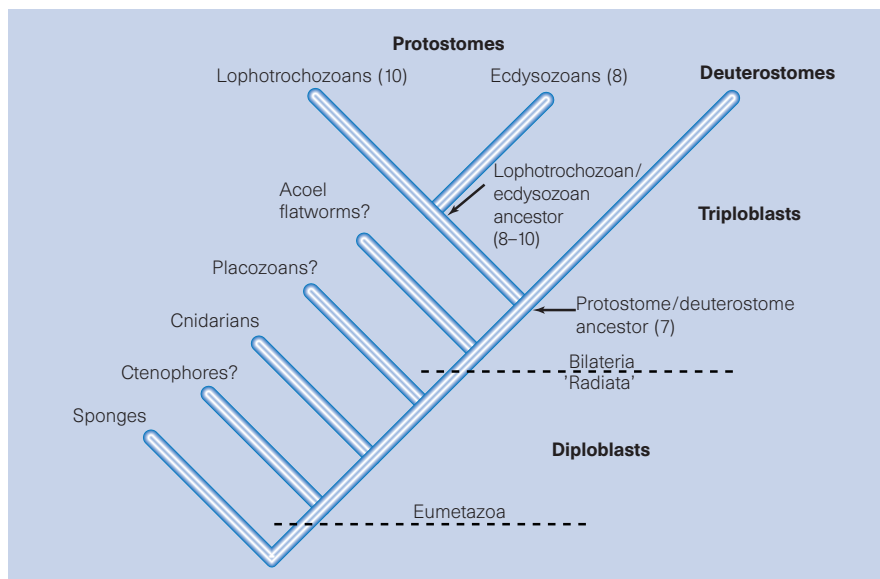


Figure 1 The main evolutionary relationships and higher-level taxonomic groupings of multicellular animals (metazoans). Important criteria that distinguish animal body plans include the presence of two (diploblast) or three (triploblast) well-defined cell layers, and the presence or absence of bilateral symmetry and an anterior–posterior axis (Bilateria or Radiata, respectively). All bilaterally symmetrical animals (bilaterians) are triploblasts. In the Bilateria, protostomes and deuterostomes are distinguished by a range of embryological characteristics. De Rosa *et al.*¹ have investigated Hox genes from the two main protostome groups — brachiopods of the lophotrochozoan clade (whose members share a similar ciliated larva), and priapulids of the ecdysozoan clade (characterized, in part, by moulting of an exoskeleton). The minimal estimate for the number of Hox genes at key points in evolution (in parentheses) suggests that Hox-gene expansion occurred early in the bilaterian lineage, before the divergence of protostomes and deuterostomes. The composition of Hox genes in the basal phyla (lower on the tree) may be crucial in determining correlations between Hox genes and new developments in body plans. Relationships between these phyla are, however, in dispute.

protostome invertebrates, in turn, consist of two main clades — the lophotrochozoans (including molluscs, annelids and brachiopods), and the ecdysozoans (including insects, crustaceans, nematodes and priapulids). These analyses threatened to rearrange almost all previously ‘accepted’ metazoan relationships based on structural characteristics, and, thus, our understanding of how and when evolutionary transitions appeared.

It is difficult to determine the origins of new phyla, because the body plans of almost all living animals appeared in a very short time, over half a billion years ago, and have remained essentially unchanged ever since. The evolutionary information from most genes (including structural RNAs) is lost by random mutation over this evolutionary timescale. It would be ideal to examine evolution of the developmental regulatory genes that are involved in generating metazoan body plans. These genes should have evolved when body plans were rapidly diversifying (so are shared by all metazoans), and their sequences have not since evolved randomly, but have paralleled the stasis of the body plans they helped to create.

In the early 1980s, developmental geneticists identified a group of genes — the Hox genes — that seemed to possess many of

these characteristics. The Hox genes encode transcription factors, and they have been found in all metazoans examined. Their function is to control the development of position-specific adult structures along the anterior–posterior axis. The Hox genes can be divided into distinct classes (called paralogue groups) found next to each other in the genome, making up a ‘cluster’. There is an argument that the evolution of more complex body plans is driven by an increase in the number of developmental regulatory genes, such as those in the Hox cluster³. Vertebrates have roughly four times as many Hox genes as their sister group, the cephalochordates (a group of invertebrate animals most closely related to the clade containing the vertebrates), whereas a lowly soil nematode has fewer than half the number of Hox genes found in the cephalochordates⁴. So, it is of interest to determine the complement of Hox genes at key nodes of metazoan evolution.

De Rosa *et al.*¹ surveyed the genomes of some key metazoan taxonomic units for Hox genes. Even though Hox genes contain a maximum of only about 75 amino acids of evolutionary information, the authors analysed them in relation to a growing database of Hox sequences. Their results support the reorganization of protostome meta-

zoans^{2,5} into the lophotrochozoans and ecdysozoans (Fig. 1). De Rosa and colleagues also found that there was a much larger — perhaps even complete — Hox cluster consisting of as many as ten Hox genes already present deep in the evolution of Bilateria. This is a minimum estimate¹, and there may have even been more genes in the cluster, early in metazoan evolution, before the divergence of protostomes and deuterostomes.

At present, though, it is not at all clear how, or whether, Hox genes were involved in the diversification of different animal phyla within the Bilateria. By studying more bilaterians, we might be able to resolve the exact origins and identities of individual paralogue groups, and the gene duplications and losses specific to particular taxonomic units. But such studies are unlikely to identify the origins of the Hox cluster itself, or give any insights into the origins of complex bilaterian body plans.

So, if the size of Hox clusters has not changed considerably within the Bilateria (with the possible exception of vertebrates), we must go even further back in metazoan evolution to determine when the Hox cluster expanded, and how this might relate to development of body plans. Although controversial in their placement, several taxonomic units (Fig. 1) may show intermediate stages of Hox-gene evolution. The cnidarians, for example (animals such as sea anemones, corals and jellyfish, which lack overt signs of bilateral symmetry), have Hox genes^{6,7}. Of course, using Hox genes as evolutionary measuring sticks to determine metazoan relationships is one thing, but understanding the evolution of their function is another. The expression patterns for some Hox genes in a handful of animals are available, but these have not been analysed with the same kind of evolutionary rigour that is now standard for molecular sequence analysis. There is no doubt that Hox genes are trying to tell us something about metazoan evolution — even if it is not what we were expecting to hear. □

Mark Q. Martindale is at the Kewalo Marine Laboratory, Pacific Biomedical Research Center, University of Hawaii, Honolulu, Hawaii 96813, USA.

e-mail: mqmartin@hawaii.edu

Matthew J. Kourakis is in the Department of Organismal Biology and Anatomy, University of Chicago, Chicago, Illinois 60637, USA.

e-mail: kour@midway.uchicago.edu

1. de Rosa, B. *et al.* *Nature* **399**, 772–776 (1999).

2. Aguinaldo, A. M. A. *et al.* *Nature* **387**, 489–493 (1997).

3. Pendleton, J. W., Nagai, B. K., Murtha, M. T. & Ruddle, F. H. *Proc. Natl Acad. Sci. USA* **90**, 6300–6304 (1993).

4. García-Fernández, J. & Holland, P. W. H. *Nature* **370**, 563–566 (1994).

5. Belavoine, G. C. *R. Acad. Sci.* **320**, 83–94 (1997).

6. Finnerty, J. R. *Curr. Top. Dev. Biol.* **40**, 212–251 (1998).

7. Finnerty, J. R. & Martindale, M. Q. *Evol. Dev.* (in the press).

Mantle geochemistry

Unmixing Hawaiian cocktails

Alex N. Halliday

Geochemical data can tell us a great deal about processes in Earth's mantle, but obtaining that data and interpreting it are far from straightforward. If confirmation of that point were needed it comes from papers by Lassiter and Hauri¹, and Putirka², both of which involve analysis of Hawaiian lavas but which come to rather different conclusions about their origins.

Lassiter and Hauri present evidence from osmium isotopes to support the idea^{3,4} that the Hawaiian mantle plume — like other plumes, a hot and buoyant jet-like structure that melts as it approaches Earth's surface — is an ancient slab of subducted ocean floor. But they go further and claim that the original variations in isotopic composition with depth in the ocean floor have survived mantle stirring and can now be seen as geographically distributed trends among the Hawaiian lavas. This is remarkable, because it implies that there was little distortion of the slab despite a billion years of mantle convection.

The magma from melting plumes under places such as Hawaii is known as ocean island basalt (OIB), and is chemically and isotopically distinct from the mid-ocean ridge basalt (MORB) of the ocean floor found at divergent plate boundaries. Their

production requires long-lived and geometrically distinct source regions (Fig. 1). On the basis of studies with lead isotopes (Box 1), it was proposed^{3,4} that OIB is produced by melting slabs of old MORB that have been recycled by subduction, the process by which ocean floor material is dragged down into the mantle at ocean trenches (not to be confused with the production of island arc magmas above subduction zones at convergent plate boundaries). In this slab-recycling model, the ocean floor material that has passed down through the subduction zone is stirred through the hot convecting mantle over hundreds of millions of years, and is then brought back to the surface in a plume to provide a source for OIB.

One snag in this scheme is that chemical data for OIB appeared to give a very different picture of how the mantle must work. For example, the trace-element ratios cerium/lead and niobium/uranium are greatly modified by ocean-floor alteration and arc processes but not by melting. So, according to the slab-recycling model, Ce/Pb and Nb/U ratios should also be relatively heterogeneous in OIB. In the 1980s, it was discovered⁵ that, except under unusual circumstances⁶, they are surprisingly uniform. Paradoxically, most erupted OIB are isotopically more heterogeneous than MORB^{6,7}, but are chemically less variable (although this could be due to differences in the way the magma reaches the surface). Average OIB should also be very different in Nb/U and Ce/Pb from average MORB, but in fact they are identical⁵. How one generates isotopic heterogeneity in OIB without affecting these trace element ratios became a dilemma.

For all that Hawaii is Earth's largest active

intra-plate volcanic centre, and a wonderful natural laboratory for Earth scientists, until recently Hawaiian magmas did not figure prominently in the slab-recycling models. Their high $^3\text{He}/^4\text{He}$ ratios⁸ do not readily fit such models, although other geochemical features do so nicely. Oxygen isotope data^{9,10} provide evidence that the Hawaiian lavas incorporated components that had interacted with seawater, and assimilation of the oceanic lithosphere through which they had erupted was considered a likely explanation¹⁰. Some of the lavas also have unusual chemical compositions consistent with the incorporation of crystal-rich components from deep oceanic crust¹¹, and this again could be interpreted to result from assimilation of oceanic lithosphere. Both features should vary with depth in the oceanic lithosphere.

The alternative, as Lassiter and Hauri¹ propose, is that these features in fact varied with depth in the original ocean floor that was subducted to produce the mantle plume itself. They provide Os data for gabbroic xenoliths exhumed from the deep oceanic lithosphere under Hawaii and show that they are too rich in ^{187}Os to be a suitable source component for the lavas. So they argue that the Os and O isotope variations come from melting recycled slabs of ancient ocean floor and associated sediments. At present the geochemical variations are distributed as parallel geographical trends among the Hawaiian lavas with a scale of several kilometres. Lassiter and Hauri contend that these variations span horizontal dimensions roughly equivalent to the depth dimensions in the original ocean floor as a result of the ocean floor being brought up to the surface 'end on' in the plume.

This conclusion raises several questions. How can slabs be so well preserved after a billion years of cycling through the mantle? The $^3\text{He}/^4\text{He}$ ratios and scale of the Hawaiian plume imply that it is derived from the lower

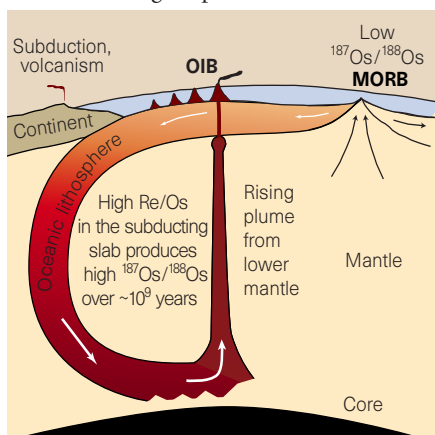


Figure 1 Production of mid-ocean ridge basalt (MORB), and of ocean island basalt (OIB) such as that which forms Hawaii, as seen in the slab-recycling model. Subduction of oceanic lithosphere continuously transports rocks that are produced near the Earth's surface, in particular MORB and associated sediments, back into the deep mantle. These introduced heterogeneities are stirred and distorted by convection in the hot mantle. Eventually they can become buoyant and rise as plumes that melt to produce OIB. The time taken for this recycling process is thought to be typically about a billion years. By the time the plume melts to produce OIB it has 'aged' isotopically and has higher $^{187}\text{Os}/^{188}\text{Os}$ than the surrounding mantle.

Box 1: Isotopic tracing of slab material

Long-lived radioactive isotopes such as ^{238}U and ^{187}Re produce changes in the isotopic compositions of their daughter elements (Pb and Os) in the mantle. The higher the parent-to-daughter ratio (Re/Os, for example), the greater the change in the isotopic composition of the daughter element in a given time. These parent/daughter ratios are altered by processes such as melting and low-

temperature interaction with water. The Re/Os ratio is also affected by assimilation of sulphides and earlier crystallized materials. Subducted slabs of altered basalt and surface sediments therefore have U/Pb and Re/Os values that differ from those in the mantle that produced the original MORB material. So over (say) a billion years the result is Pb and Os isotopic compositions that differ

from those in modern MORB (Fig. 1). The Pb isotopic compositions of OIB are highly variable and consistent with time-integrated high U/Pb, confirming the basic model that OIB is derived from old recycled slabs^{3,4}. It is only recently that the Re/Os technique, as used by Lassiter and Hauri¹, has been refined to the extent that it could be used to test the earlier results.

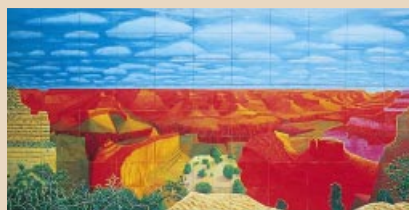
A.N.H.

Geomorphology

A bigger Hockney

For some weeks London tube-travellers were treated to posters of this David Hockney painting, "A Closer Grand Canyon, 1998", advertising the annual summer exhibition of the Royal Academy of Arts at Burlington House, Piccadilly, in central London. The posters, let alone this picture, scarcely do justice to the scale of the work, which is painted in oils on 96 canvases, with an overall dimension of about 11 by 24 feet. A room at the exhibition is devoted to six such vivid depictions of the canyon.

According to Hockney, it is the space rather than geology of the canyon he has tried to capture in this and similar studies: "the sense of the eye meandering through a vast landscape". Photography failed to do so. So pastel drawings were used as the basis for the final works in oils, which have



come to London after being on display at the Centre Georges Pompidou in Paris.

The summer exhibition includes art of various forms, including architectural models, both by academicians such as Hockney and non-academicians, and runs until 15 August. Hockney's Grand Canyon pastels and other drawings can be seen at Annely Juda Fine Art, 23 Dering Street, London W1, in a separate showing which starts on 30 June and ends on 18 September.

Tim Lincoln

mantle, so why do these heterogeneities not get smeared out instead of coming back into the uppermost mantle with roughly the same dimensions as originally formed on the ocean floor? Furthermore, why do the incompatible trace-element ratios of Hawaiian magmas not look more like those of altered oceanic basalt? If the Hawaiian magmas are re-melting well-preserved ancient ocean floor in a plume, why are the trace elements not highly variable as in modern altered oceanic crust? Some of these more mobile elements may be partially removed in the subduction zone, but there should be some trace of these processes in the Nb/U and Ce/Pb ratios for example.

Some of these issues are addressed by Putirka², who uses trace-element compositions of lavas combined with melting models to 'retrieve' the original composition of the mantle beneath Hawaii. He argues that the degree of variability in mineralogy and chemical composition in the source is in fact very small. Subducted slabs should be highly heterogeneous in major and trace elements, and in mineralogy, not just because they have been altered by seawater but because the upper portion of the slab is metamorphosed basaltic rocks with high levels of sodium, calcium and aluminium (very distinct from the more magnesium-rich and Al-depleted peridotite that makes up the bulk of the mantle). The exact mineralogy would depend on the pressure, but rocks called garnet pyroxenite and eclogite are the most likely types. The mineralogy also affects the partitioning of certain trace elements. Furthermore, even if the mineralogy remains the same, the degree to which key trace elements partition between melt and silicate minerals is pressure dependent.

Putirka constructs models of partition-

ing between peridotite and melt as a function of temperature and pressure (depth). The sodium/titanium ratio is particularly sensitive to pressure during melting, and from such data Putirka argues that the Hawaiian mantle does not vary greatly in major-element composition. If recycled ocean floor was involved, a reasonably efficient homogenization of any such chemical heterogeneity must have occurred. Furthermore, Putirka claims that the isotopic heterogeneities, such as those used by Lassiter and Hauri¹, are related to the depth of melting of the mantle and are not simply lateral variations in a plume.

Putirka's arguments run counter to those of Hauri¹², who demonstrated correlations between the isotopic and major-element compositions of Hawaiian lavas that he

interpreted to reflect the incorporation of up to 20% of silica-rich 'dacitic' melts from recycled oceanic crust (now converted to garnet pyroxenite). Putirka's arguments are based more on trace elements and their partitioning, and he admits that not all of the variability in Hawaii can be explained by a homogeneous source. Nonetheless, he disputes Hauri's claim that a garnet pyroxenite (recycled slab) component is so important.

Where does all this leave us? We know that subduction occurs and that ultimately subducted slabs must contribute to mantle heterogeneity as evident in OIB. But which isotopic features are the product of re-melting of ancient subducted components? One might get cynical about mantle geochemistry because basic questions such as these keep recurring. Particularly with integrated chemical and isotopic studies, however, they look more tractable than they ever have been. We must keep going. □

Alex N. Halliday is in the Institute for Isotope Geology and Mineral Resources, Department of Earth Sciences, ETH Zentrum, NO C61, Sonneggstrasse 5, CH-8092 Zürich, Switzerland. e-mail: halliday@erdw.ethz.ch

1. Lassiter, J. C. & Hauri, E. H. *Earth Planet. Sci. Lett.* **164**, 483–496 (1998).
2. Putirka, K. J. *Geophys. Res.* **104**, 2817–2829 (1999).
3. Hofmann, A. W. & White, W. M. *Earth Planet. Sci. Lett.* **57**, 421–436 (1982).
4. Chase, C. G. *Earth Planet. Sci. Lett.* **52**, 277–284 (1981).
5. Hofmann, A. W., Jochum, K. P., Seufert, M. & White, W. M. *Earth Planet. Sci. Lett.* **79**, 33–45 (1986).
6. Halliday, A. N. et al. *Earth Planet. Sci. Lett.* **133**, 379–395 (1995).
7. Allegre, C. J., Schiano, P. & Lewin, P. *Earth Planet. Sci. Lett.* **129**, 1–12 (1995).
8. Kurz, M. D., Jenkins, W. J., Hart, S. R. & Clague, D. *Earth Planet. Sci. Lett.* **66**, 388–406 (1983).
9. Garcia, M. O., Jorgenson, B. A., Mahoney, J. J., Ito, E. & Irving, A. J. *J. Geophys. Res.* **98**, 537–550 (1993).
10. Eiler, J. M., Valley, J. W. & Stolper, E. M. *J. Geophys. Res.* **101**, 11807–11813 (1996).
11. Hofmann, A. W. & Jochum, K. P. *J. Geophys. Res.* **101**, 11831–11839 (1996).
12. Hauri, E. H. *Nature* **382**, 415–419 (1996).

Cancer

DNA damage enables p73

Eileen White and Carol Prives

The p53 gene product is a critical human tumour suppressor. In response to cellular stresses such as DNA damage and oxygen starvation, p53, a sequence-specific transcriptional activator, induces cell-cycle arrest or programmed cell death (apoptosis)¹.

Until 1997, scientists working on p53 assumed that this gene was unique. So, the discovery of two p53-related genes — p73 and p63, each of which is comprised of several isoforms — was a big surprise². As might be predicted, both genes encode proteins with transactivation, DNA-binding and tetramerization domains, and they share considerable homology with p53. Some iso-

forms of p63 and p73 can transactivate many of p53's targets (albeit to differing extents), and some forms also induce cell-cycle arrest and apoptosis. Hence the idea that certain cellular responses previously assumed to be 'p53 independent' might, in fact, be attributable to these relatives of p53.

One small comfort for p53 researchers was the possibility that only p53 can be induced when cells are exposed to stress signals such as DNA damage. Now, this assumption has been challenged too. Papers by Gong et al.³, Agami et al.⁴ and Yuan et al.⁵ (pages 806, 809 and 814 of this issue) together provide evidence that p73 is a target of the non-receptor tyrosine kinase c-Abl in response to

DNA damage. Thus another functional parallel emerges between p53 and p73.

The c-Abl tyrosine kinase is the cellular homologue of v-Abl — the product of the oncogene carried by Abelson murine leukaemia virus⁶. Expression of v-abl induces lymphomas in mice, and activation of c-abl by chromosomal translocation is associated with chronic myelogenous leukaemia and some forms of acute lymphoma in humans. Furthermore, c-Abl deficiency causes diverse developmental effects in the mouse. So, c-Abl gain-of-function promotes cancer, whereas c-Abl loss-of-function has profound developmental consequences.

Insight into how c-Abl acts came from the identification of its upstream regulatory signals and downstream targets. DNA damage (from ionizing radiation and alkylating agents) activates c-Abl's kinase activity (Fig. 1), providing a connection between c-Abl and the response to DNA damage⁷. The absence of c-Abl disrupts the G₁-S checkpoint in response to DNA damage. In cells that receive DNA damage, the c-Abl protein is phosphorylated, suggesting that there is another protein kinase upstream in the signalling pathway. Indeed, the ataxia-telangiectasia-mutated (ATM) gene product, a component of the DNA-damage checkpoint⁸, interacts with and phosphorylates c-Abl^{9–11}. The ATM protein is a widely expressed member of the family of protein kinases with similarities to phosphatidylinositol 3-kinases. It is activated by DNA damage, and loss of ATM function in human and mouse cells causes defects in DNA repair and cell-cycle checkpoint control. The result is radioresistant DNA synthesis and, not sur-

prisingly, humans and mice with compromised ATM function are prone to cancer.

There is strong evidence that a key downstream target of ATM is p53, which is phosphorylated and stabilized by ATM in response to DNA damage^{12–14}. Although functional interactions between c-Abl and p53 have been observed, they are likely to be indirect because p53 is not phosphorylated on tyrosine. Activation of c-Abl and p53 by ATM probably implements distinct pathways, and we now have an indication that p73 may be a downstream effector of c-Abl.

The three groups report that p73 is affected by some forms of DNA damage. Gong *et al.*³ demonstrate that levels of p73 protein are increased after cells are treated with cisplatin. Agami *et al.*⁴ and Yuan *et al.*⁵ show that p73 is phosphorylated after γ -irradiation of cells and by c-Abl. Activation of p73 requires functional c-Abl — neither phosphorylation nor stabilization occur in c-Abl-deficient cells^{3–5}. It is not clear why p73 stabilization is seen with cisplatin, but not with γ -irradiation or ultraviolet light, nor why tyrosine phosphorylation of p73 is observed with γ -irradiation but not with cisplatin. Nonetheless, the effect of c-Abl on p73 is likely to be direct, because Agami *et al.* and Yuan *et al.* provide evidence that c-Abl's Src-homology-3 (SH3) domain interacts with p73's carboxy-terminal PXXP motif. All three groups observe that the proapoptotic activity of p73 is potentiated by c-Abl, and diminished in c-Abl null cells, consistent with a model in which specific DNA-damage signals are channelled through c-Abl and thence to p73 (Fig. 1). If this is correct, p73-deficient cells should have defective DNA-damage checkpoint control.

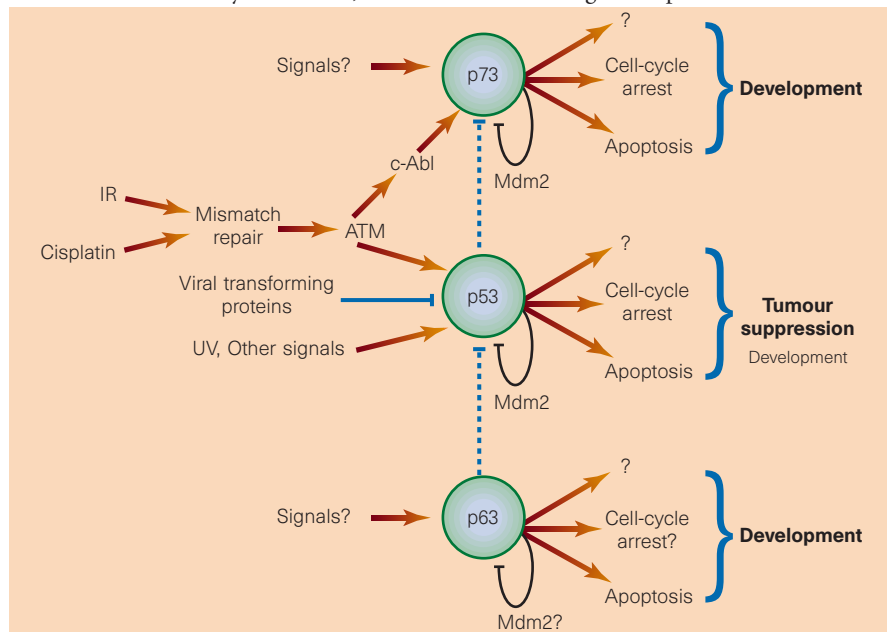


Figure 1 Regulatory targets and downstream signals affecting p53 and the related proteins p63 and p73. Gong *et al.*³, Agami *et al.*⁴ and Yuan *et al.*⁵ have shown that, like p53, the p73 protein is induced in response to DNA damage, via the non-receptor tyrosine kinase c-Abl. Dotted lines indicate inhibition of p73 by mutant forms of p53, and inhibition of p53 by alternately spliced forms of p63. IR, ionizing radiation; ATM, ataxia-telangiectasia-mutated protein; UV, ultraviolet light.



100 YEARS AGO

Prof. Arrhenius contributes to the *Revue Générale des Sciences* an interesting account of his investigations into the causes of secular variations of temperature at the earth's surface. It is shown that widespread changes of mean temperature are more likely to be due to variations in the proportion of terrestrial rays absorbed by the atmosphere than to any variation connected with the solar rays, and that the absorption of terrestrial rays is most likely to be affected by changes in the amount of carbonic acid present in the atmosphere. Using Langley's data, it is calculated that if the amount of carbonic acid were diminished by a little more than half, the temperature would be lowered by about 4°·5 C., while an increase to two and half or three times the present amount would raise the temperature about 8°·5C., corresponding to the conditions of Glacial and Eocene times respectively.

From *Nature* 22 June 1899.

50 YEARS AGO

During the Academy celebrations in 1945, I had asked to see Lysenko and his results, but had been told that he was too busy. However, after repeated requests, it was suddenly announced that he would lecture next day, and I went to listen, accompanied by Prof. Ashby, and by an excellent interpreter who was also a biologist. Her running translation was at one moment drowned by a burst of laughter from the audience. On my asking her afterwards what had provoked this, it appeared that Lysenko was discussing Mendelian dominance and segregation, which his opponents sometimes brought up against him. Dominance, he said, was easy to explain on his own theories; it was the 'assimilation' ... of one heredity by a second, after a cross. But segregation (of recessive characters in F_2)? That also was easy. 'We know in our persons,' he said, 'that assimilation (or digestion) is not always complete. When that is so, what happens? We belch. Segregation is Nature's belching; unassimilated hereditary material is belched out' (presumably in a 1 : 3 ratio!). I cite this remark as further evidence of the fact that, scientifically, Lysenko can only be described as illiterate. Julian Huxley

From *Nature* 25 June 1949.

As the boundaries continue to blur between p53 and its relatives, compelling differences remain. For example, despite their similarities, the three genes seem to have very different functions. Whereas deletion of *p63*^{15,16} and *p73* (F. McKeon, personal communication) has dramatic developmental consequences in mice, *p53* null mice develop normally¹ (with some interesting exceptions). But *p53* null mice are highly prone to tumours. That *p53* is unique in serving as a tumour suppressor is supported by the fact that, in human cancers, loss or mutation of *p73* or *p63* seems to be infrequent². Additionally, only *p53* is susceptible to inactivation by the SV40 T antigen, the adenovirus E1B 55K protein and the human papillomavirus E6 protein. Finally, although Mdm2 binds to *p53* and *p73* (Fig. 1), only *p53* is degraded as a result of this interaction².

Why might *p53* alone be a tumour suppressor? There are several possible answers, all supported by published reports². First, the tissue distribution of *p63* and *p73* is more restricted than that of *p53*. Second, the downstream targets of *p53* and its relatives may differ, and perhaps *p53* is more effective in inducing key apoptosis-related targets under some circumstances. Third, there may be a broader range of upstream regulators that can signal to *p53*. And fourth, interactions between some forms of *p53* and its relatives have been documented. For instance, tumour-derived *p53* mutants can repress transactivation and apoptosis induced by *p73*, and amino-terminally-truncated forms of *p63* can repress wild-type *p53* (ref. 2). So, cross-talk between *p53* and its relatives may contribute to their different roles in tumorigenesis.

It has been suggested that *p53* is the most recently evolved member of this family. We might speculate that, with the development of more complex multicellular organisms, a need arose for a more versatile stress-response factor — one that can respond to and transmit a more diverse set of signals to a more complex set of targets. Whatever the explanation for the differences among *p53* family members, there is much work to be done to find out how these genes regulate important processes in cells, and why only *p53* suppresses tumour formation. □

Eileen White is at the Howard Hughes Medical Institute, Center for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry, Cancer Institute of New Jersey, and Rutgers University, 679 Hoes Lane, Piscataway, New Jersey 08854, USA.

e-mail: ewhite@mbcl.rutgers.edu

Carol Prives is in the Department of Biological Sciences, Columbia University, Sherman Fairchild Center for the Life Sciences, 816 Fairchild Building, New York, New York 10027, USA.

e-mail: clp3@columbia.edu

1. Ko, L. J. & Prives, C. *Genes Dev.* **10**, 1054–1072 (1996).
2. Kaelin, W. G. Jr *Oncogene* (in the press).

3. Gong, J. *et al. Nature* **399**, 806–809 (1999).
4. Agami, R., Blandino, G., Oren, M. & Shaul, Y. *Nature* **399**, 809–813 (1999).
5. Yuan, Z.-M. *et al. Nature* **399**, 814–817 (1999).
6. Van Etten, R. A. *Trends Cell Biol.* **9**, 179–186 (1999).
7. Kharbanda, S. *et al. Nature* **376**, 785–788 (1995).
8. Rotman, G. & Shiloh, Y. *Hum. Mol. Genet.* **7**, 1555–1563 (1998).

9. Baskaran, R. *et al. Nature* **387**, 516–519 (1997).
10. Shafman, T. *et al. Nature* **387**, 520–523 (1997).
11. Yuan, Z.-M. *et al. Nature* **382**, 272–274 (1996).
12. Banin, S. *et al. Science* **281**, 1674–1677 (1998).
13. Canman, C. E. *et al. Science* **281**, 1677–1679 (1998).
14. Khanna, K. K. *et al. Nature Genet.* **20**, 398–400 (1998).
15. Mills, A. A. *et al. Nature* **398**, 708–713 (1999).
16. Yang, A. *et al. Nature* **398**, 714–718 (1999).

Evolutionary genetics

No sex please, we're fungi

Ian R. Sanders

According to evolutionary theory, the advantage of sex is that recombination shuffles together new combinations of genes, thereby producing genetic variation and allowing deleterious mutations to be purged^{1–3}. But some extremely successful organisms are both asexual and ancient⁴: the very existence of such 'scandalous' asexuals⁵ flies in the face of theory.

Among them are counted the arbuscular mycorrhizal fungi, the Glomales, studies of the genome structure and organization of which now reveal some remarkable phenomena. The papers concerned appear in *Fungal Genetics and Biology*⁶ and *Gene*⁷. Most notably, they identify a striking degree of divergence in the ribosomal sequences of individuals among the Glomales.

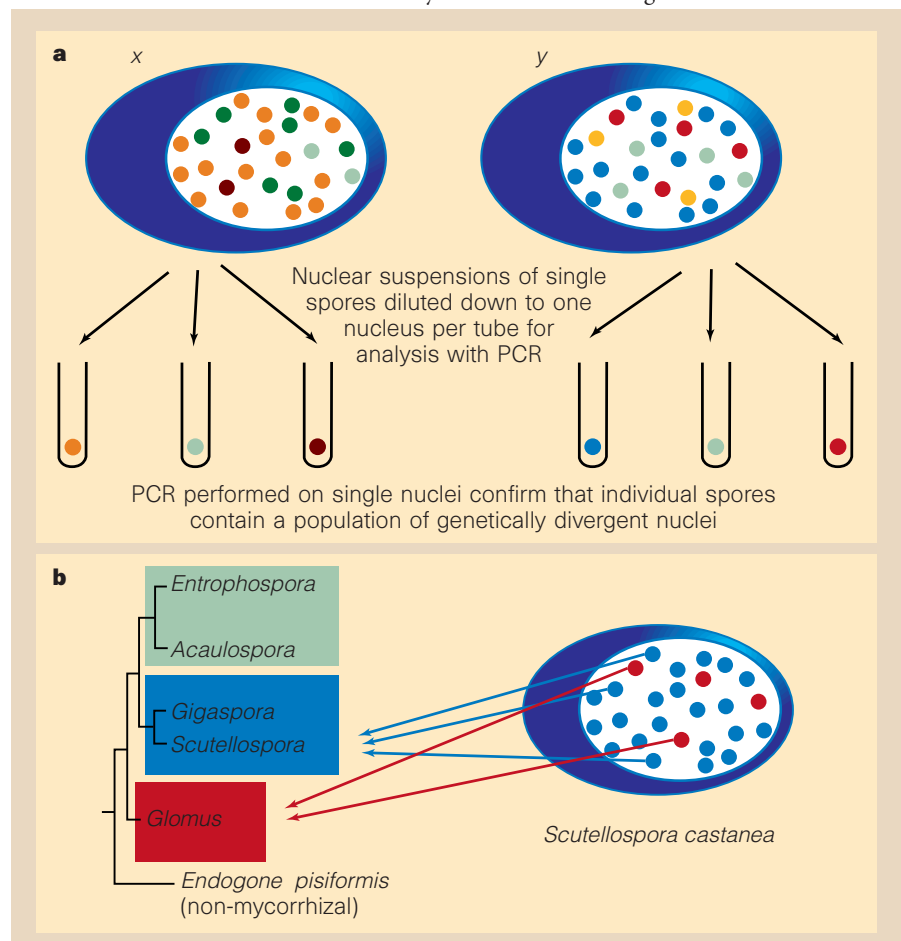


Figure 1 Genetics of asexual mycorrhizal fungi of the order Glomales. a, Hijri *et al.*⁶ demonstrate that different spores of *Scutellospora castanea* (x and y) contain different sequences of ribosomal DNA, but that each spore does not have the same complement of sequences. Further, they show that rDNA differs between nuclei from the same fungal individual, which may be because lack of recombination has allowed the multiple copies of rDNA to diverge. b, Hosny *et al.*⁷ find that different sequences of the 18S gene from an isolate of *S. castanea* group into two different Glomales genera, *Scutellospora* and *Glomus*. The Glomales consists of three different families: Glomaceae (red), Gigasporaceae (light blue) and Acaulosporaceae (green), and the first two are thought to have diverged 353–367 million years ago¹⁵. The discovery of rDNA sequences in *S. castanea* that match genera in two different families suggests that genetic diversity in these fungi may have been increased by acquisition of DNA from their ancestors.

These fungi form symbioses — mycorrhizas — with plant roots which help plants to acquire mineral nutrients from the soil and also determine plant biodiversity and ecosystem function⁸. They are truly ancient, having remained largely morphologically unchanged since plants first colonized the land around 400 million years ago⁹. No sexual stage in the Glomales life cycle has been observed.

Polymorphism in the ribosomal DNA encoding the internal transcribed spacer (ITS), and the 5.8S and 18S subunits, occurs inside individual spores of fungi of the genus *Glomus*^{10,11}. This is unusual because it is widely thought that sequences of multiple copies of ribosomal DNA are kept the same by a process known as concerted evolution¹², in which the joint evolution of two or more related genes occurs as if they constitute a single locus. It is on the assumption of concerted evolution that the ribosomal sequences are so widely used for taxonomy and phylogenetics¹³. The existence of different sequences of ribosomal genes in *Glomus* implies that recombination does not occur, because the mechanisms that keep the sequences of rDNA copies the same operate most frequently during recombination events.

Both of the new studies^{6,7} looked at

another fungus within the Glomales, *Scutellospora castanea*, and worked with exactly the same isolate¹⁴ which was maintained in culture in the roots of plants. Using the polymerase chain reaction (PCR), six different ITS sequences and 13 of the 18S gene were picked out, respectively, from genomic DNA and from clones in an *S. castanea* genomic DNA library. This does not mean that variation is limited to six different sequences of each. Hijri *et al.*⁶ confirmed that several different ITS sequences (including those of the 5.8S gene) occur in a single spore, but that each spore of this isolate does not have the same complement of the different sequences (Fig. 1a).

The Glomales are coenocytes — that is, many nuclei are enclosed within one cell wall — and it could be that different rDNA sequences occur on different nuclei in which the rDNA sequences have diverged due to a lack of recombination. To test this possibility, Hijri *et al.*⁶ diluted nuclear suspensions from single spores to an expected one-nucleus-per-sample and performed PCR with specific primers that amplify the different rDNA sequences. Their results indicate that, indeed, different rDNA sequences are carried on different nuclei (Fig. 1a); so it seems that an individual of these fungi is, in genetic terms,

actually a population of discrete nuclei.

Perhaps more astounding, however, is just how divergent the ribosomal gene sequences are. Hosny *et al.*⁷ carried out a phylogenetic analysis of 18S sequences from *S. castanea* and from species of each of the other Glomales genera. Most sequences grouped in the genus *Scutellospora*, but two others from the same *S. castanea* isolate were so divergent that they clustered in the genus *Glomus* (Fig. 1b).

A contaminant in the genomic library, perhaps? It seems not, because the *Glomus*-like rDNA sequences were successfully amplified again from single spores of this isolate of *S. castanea*. According to previous phylogenetic analyses of the Glomales, the family known as Glomaceae, which contains the genus *Glomus*, diverged from the other mycorrhizal fungi that were later to form the Acaulosporaceae and the Gigasporaceae (containing the genus *Scutellospora*), between 353 and 367 million years ago¹⁵.

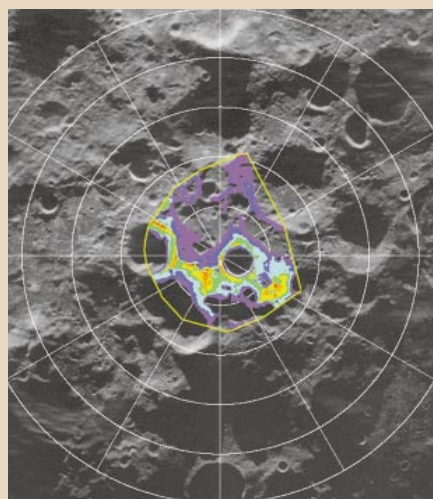
How then can we explain the observations reported in these two studies? The results support the idea that genetic drift — the random change of allelic frequencies in a population — has occurred in the absence of recombination, leading to genetically divergent nuclei. But at the same time it seems

Lunar exploration

Polar endeavours

Almost exactly 85 years ago, Sir Ernest Shackleton set out to cross the Antarctic continent. After his ship was frozen into the ice pack, the expedition waited and worked through the long polar nights, then the long polar days, before Shackleton managed to get them all rescued, a story that continues to generate popular books¹. In honour of his exploits, Shackleton had a lunar crater (just to the right of centre in this image of the lunar south pole) named after him. That crater is now getting attention as a part of another story that involves ice, long days and nights — and perhaps, ultimately, human exploration.

Lunar scientists have long known that because the Moon's axis of rotation is almost perpendicular to its path around the Sun, the polar regions could have high points that are permanently sunlit, and crater floors that are permanently in shadow. This false-colour image (inside the yellow line, and overlaid on a radar image of the region) shows the percentage of the lunar day for which a given location is illuminated (white, orange and red receive the most illumination, the clear regions receive none), based on a newly reported analysis of images taken in 1994 by the Clementine spacecraft². A separate radar study³ has confirmed that many regions,



including the bottom of Shackleton Crater, are likely to be permanently shadowed. Conversely, some regions are sunlit most of the time and would be prime locations for lunar bases, or at least for solar arrays to support a polar base. For example, the white region at the left (poleward) edge of Shackleton Crater is sunlit more than 80% of the time, and there is a ridge 10 km away that is sunlit 90% of the time that the first spot is not.

A polar base would be more attractive if the permanently shadowed crater floors have managed to cold-trap water ice, which

could be mined for life support or fuel.

Data from the neutron spectrometer aboard the Lunar Prospector spacecraft⁴ apparently showed abundant hydrogen at the poles, possibly, though not definitely, in the form of water ice; similarly, a Clementine radar experiment may⁵ or may not⁶ have provided evidence for ice.

At the end of July, with its funding expired and its fuel nearly gone, Lunar Prospector will be crashed into a crater near the pole (the larger crater above and to the right of Shackleton) and Earth-bound telescopes will search for evidence of water released by the impact⁷. The chances of the experiment detecting water are no better than 10%, but this is just the next step in a search for the conditions that might make exploration of the lunar poles possible.

Timothy D. Swindle

Timothy D. Swindle is in the Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721-0092, USA.

e-mail: tswindle@U.Arizona.EDU

1. Alexander, C. *The Endurance: Shackleton's Legendary Antarctic Expedition* (Knopf, New York, 1998).

2. Bussey, D. B. J., Spudis, P. D. & Robinson, M. *Geophys. Res. Lett.* **26**, 1187–1190 (1999).

3. Margot, J. *et al. Science* **284**, 1658–1660 (1999).

4. Feldman, W. C. *et al. Science* **281**, 1496–1500 (1998).

5. Nozette, S. *et al. Science* **274**, 1495–1498 (1996).

6. Simpson, R. A. & Tyler, G. L. *J. Geophys. Res.* **104**, 3845–3862 (1999).

7. Goldstein, D. B. *et al. Geophys. Res. Lett.* **26**, 1652–1656 (1999).

unlikely that the nuclei have also harboured rDNA sequences that have then remained conserved for 400 million years. Probably the most likely way to explain their occurrence is hyphal anastomosis (cross-connection), and the exchange of nuclei. Clearly, more information is now needed on DNA polymorphisms within and among spores of *Glomales* species from regions of DNA other than the multi-copy ribosomal genes. One possibility is that infrequent recombination occurs among nuclei. Using such sequences as markers in population-genetic models would enable a test of whether nuclei from a single spore form a recombinant population.

Meantime we are left with the distinct possibility that individuals in the *Glomales* contain a population of highly divergent nuclei that are subject to the accumulation of mutations and additional genetic material from distant sources. How do these fungi cope with the accumulation of deleterious mutations? How can they perform and regulate their cellular, developmental and metabolic processes? Phylogeneticists are faced with equally fundamental questions. Is concerted evolution in fact in operation? Is it valid to apply phylogenetic techniques to these organisms at all?

The evolution of genomes in the absence of recombination is not yet understood. But perhaps the way that these asexual fungi have

existed for 400 million years is in some way connected to the maintenance of high genetic diversity in their populations and their ability to have exchanged genetic material with their ancestors. What is clear is that evolutionary biologists and fungal geneticists are faced with a happily disconcerting puzzle. □

Ian R. Sanders is at the Botanisches Institut, Universität Basel, Hebelstrasse 1, 4056 Basel, Switzerland.

email: sanders@ubaclu.unibas.ch

1. Stearns, S. C. (ed.) *The Evolution of Sex and its Consequences* (Birkhäuser, Basel, 1987).
2. Fain, H. D. *Evol. Theory* **11**, 15–29 (1995).
3. Lynch, M., Butcher, R. B. D. & Gabriel, W. J. *Hered.* **84**, 339–344 (1993).
4. Judson, O. P. & Normark, B. B. *Trends Ecol. Evol.* **11**, 41–46 (1996).
5. Maynard-Smith, J. *Nature* **324**, 300–301 (1986).
6. Hijri, M., Hosny, M., van Tuinen, D. & Dulieu, H. *Fungal Genet. Biol.* **26**, 141–151 (1999).
7. Hosny, M., Hijri, M., Passerieux, E. & Dulieu, H. *Gene* **226**, 61–71 (1999).
8. van der Heijden, M. G. A. *et al.* *Nature* **396**, 69–72 (1998).
9. Remy, W., Taylor, T. N., Hass, H. & Herp, H. *Proc. Natl Acad. Sci. USA* **91**, 11841–11843 (1995).
10. Sanders, I. R., Alt, M., Groppe, K., Boller, T. & Wiemken, A. *New Phytol.* **130**, 419–427 (1995).
11. Lloyd-McGill, S. A. *et al.* *New Phytol.* **133**, 103–111 (1996).
12. Hoelzel, A. R. & Dover, G. A. *Molecular Genetic Ecology* (IRL, Oxford, 1991).
13. Winker, S. & Woese, C. R. *System. Appl. Microbiol.* **14**, 305–310 (1991).
14. Banque Européenne des Glomales (BEG) <http://www.bio.ukc.ac.uk/beg/>
15. Simon, L., Bousquet, J., Lévesque, R. C. & Lalonde, M. *Nature* **363**, 67–69 (1993).

Alzheimer's disease

Pinning down phosphorylated tau

Michel Goedert

Defining characteristic of several neurodegenerative diseases, including Alzheimer's disease and Pick's disease, is the formation of filamentous deposits of a microtubule-associated protein called tau in an abnormally hyperphosphorylated form. The discovery of mutations in the *tau* gene in a condition known as 'familial frontotemporal dementia and parkinsonism linked to chromosome 17' has renewed interest in the mechanisms by which dysfunction of tau causes neurodegeneration.

On page 784 of this issue, Lu *et al.*¹ describe an intriguing interaction between phosphorylated tau and a prolyl isomerase, Pin1. Prolyl isomerases enhance the rate of *cis* to *trans* isomerization of the peptide bond on the amino-terminal side of proline. Pin1 is an essential nuclear protein belonging to the parvulin family of prolyl isomerases². This group is distinct from two other prolyl isomerase families, the cyclophilins and the FK506-binding proteins, which are targets of the immunosuppressive drugs cyclosporine and FK506, respectively. Pin1 consists of a carboxy-terminal catalytic domain, as well as an amino-terminal protein-protein interaction region called a WW domain

that specifically recognizes phosphorylated serine or threonine residues preceding a proline residue (the S/T-P motif)^{2,3}.

In its short history, Pin1 has generated much interest because it regulates entry and progression through mitosis. It does this by interacting with a large set of mitosis-specific phosphoproteins, most of which can be detected by a monoclonal antibody called MPM2 (ref. 3). Lu *et al.*¹ started off armed with the knowledge that MPM2 also recognizes hyperphosphorylated tau in the brains of people with Alzheimer's disease^{4,5}. Moreover, during mitosis, tau is phosphorylated at a number of the S/T-P sites that are hyperphosphorylated in Alzheimer's^{6,7}. This prompted the authors to examine whether phosphorylated tau interacts with Pin1. And they found that tau (phosphorylated either by a mitotic cell extract or by Cdc2 kinase) does, indeed, bind to the WW domain of Pin1.

The longest isoform of tau in the human brain has 17 S/T-P sites. Of these, Lu and colleagues found that only one — phosphorylated threonine 231 (T231) — was required for the interaction with Pin1. This residue is located upstream of the microtubule-

binding repeats in a proline-rich region that is required for full activity of tau. The T231 residue is hyperphosphorylated in Alzheimer's disease and is also phosphorylated, to a certain extent, in the normal brain^{8,9}. This residue can also be phosphorylated by glycogen synthase kinase-3 β , but only after phosphorylation of serine 235 by cyclin-dependent kinase-5 or mitogen-activated protein kinase^{10,11}.

Lu and colleagues went on to gather more evidence for the interaction between tau and Pin1. First, using a pull-down assay, they showed that Pin1 binds to hyperphosphorylated tau from the brains of people with Alzheimer's disease, but not to tau from age-matched healthy brains. Tau from normal brain is notorious for the speed with which it is dephosphorylated after death¹², so it may be premature to conclude that Pin1 interacts only with hyperphosphorylated (pathological) tau.

Second, by immunoblotting, the authors detected endogenous Pin1 in the paired helical filaments (PHFs) from diseased brains. (The PHFs, which are composed of hyperphosphorylated tau, make up the pathological neurofibrillary tangles of Alzheimer's disease.) Third, using immunohistochemistry, Lu *et al.* found that recombinant Pin1 binds to pathological tau. Finally, the authors looked at localization of Pin1. In control brains they observed nuclear staining for endogenous Pin1, consistent with its known localization in non-neuronal cells. But in the brains of people with Alzheimer's disease, Pin1 staining was also associated with pathological tau in neuronal cells (although it is not clear what percentage of the tau-positive structures was also immunoreactive for Pin1).

The tau protein in PHFs from the brains of patients with Alzheimer's disease is phosphorylated at more than 20 residues, many (but not all) of which are S/T-P sites¹³. In healthy brains, tau is heterogeneously phosphorylated at between eight and ten of these residues⁹. Because of its abnormal hyperphosphorylation, tau from PHFs cannot bind to microtubules or promote microtubule assembly. Hyperphosphorylation of tau is believed to be an early event that precedes assembly into PHFs. Yet there is no experimental evidence linking hyperphosphorylation of tau to PHF assembly¹⁴ — synthetic, paired helical-like filaments can, in fact, be produced in a phosphorylation-independent way.

To examine the functional effects of the interaction between Pin1 (Fig. 1, overleaf) and tau phosphorylated at T231, Lu *et al.* used recombinant tau phosphorylated by Cdc2 kinase, with or without added Pin1. As expected, phosphorylated tau could neither bind microtubules nor promote microtubule assembly properly. But in the presence of Pin1, biological activity of the phos-

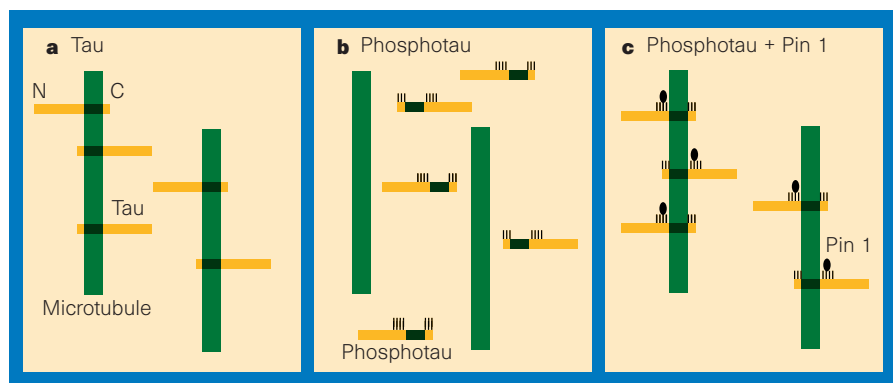


Figure 1 Pin1 restores the ability of phosphorylated tau to bind to microtubules. **a**, Non-phosphorylated tau protein (yellow) binds to microtubules (green). **b**, Cdc2 kinase phosphorylates tau at a number of sites where serine or threonine precedes a proline residue (S/T-P sites), located on either side of the microtubule-binding repeats (shown in black). As a result, tau has a reduced ability to bind to microtubules. **c**, The work of Lu *et al.*¹ indicates that Pin1 (ovals) binds to phosphorylated tau protein and restores its ability to bind to microtubules. Pin1, by itself, does not bind to microtubules. Phosphothreonine 231 in tau is both necessary and sufficient for the interaction with Pin1. The isomerase activity of Pin1 is required to restore the biological activity of phosphorylated tau.

phorylated tau was restored, an effect that seemed to depend on Pin1's prolyl isomerase activity. This surprising result indicates that isomerization of the peptide bond on the amino-terminal side of proline residue 232 in tau is enough to confer biological activity upon Cdc2-phosphorylated tau. Moreover, it suggests that the same may be true of the hyperphosphorylated tau in Alzheimer's disease. However, Cdc2 kinase phosphorylates only some of the residues that are hyperphosphorylated in the disease, and it remains to be seen whether Pin1 could restore the biological activity of tau *in vivo*.

Lu *et al.* next compared the levels of soluble Pin1 in the brains of Alzheimer's patients with those in age-matched control brains. By immunoblotting and immunoprecipitation, they saw greatly reduced levels of soluble Pin1 in the diseased brain. The most striking abnormality of tau in Alzheimer's disease is its assembly into filaments, and these results indicate that tau filaments may sequester soluble Pin1. Depletion of Pin1 in non-neuronal cells leads to mitotic arrest², so the observed depletion may have been related to ectopic expression of some cell-cycle markers in the brains of people with Alzheimer's disease¹⁵⁻¹⁷. The significance of this observation for the neurodegenerative process is unclear, although it has been suggested that aberrant expression of cell-cycle markers may cause nerve cell death by apoptosis.

The identification of mutations in the *tau* gene should allow us to study these questions more rigorously. Using such mutants, it has already been shown that a reduced ability of tau to interact with microtubules may precede hyperphosphorylation and assembly into filaments¹⁴. So, compounds that reduce the pool of functionally impaired tau could be of therapeutic value. The discovery of the interaction between Pin1 and phosphorylat-

ed tau may open the way to developing such compounds. □

Michel Goedert is at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

e-mail: mg@mrc-lmb.cam.ac.uk

1. Lu, P.-J., Wulf, G., Zhou, X. Z., Davies, P. & Lu, K. P. *Nature* **399**, 784-788 (1999).
2. Lu, K. P., Hanes, S. D. & Hunter, T. *Nature* **380**, 544-547 (1996).
3. Yaffe, M. B. *et al. Science* **278**, 1957-1960 (1997).
4. Vincent, I., Rosado, P. & Davies, P. *J. Cell Biol.* **132**, 413-425 (1996).
5. Kondratieck, C. M. & Vandr , D. D. *J. Neurochem.* **67**, 2405-2416 (1996).
6. Pope, W. B. *et al. Exp. Neurol.* **126**, 185-194 (1994).
7. Illenberger, S. *et al. Mol. Biol. Cell* **9**, 1495-1512 (1998).
8. Hasegawa, M. *et al. J. Biol. Chem.* **267**, 17047-17054 (1992).
9. Watanabe, A. *et al. J. Biol. Chem.* **268**, 25712-25717 (1993).
10. Ishiguro, K. *et al. FEBS Lett.* **325**, 167-172 (1993).
11. Goedert, M. *et al. Biochem. J.* **301**, 871-877 (1994).
12. Matsuo, E. S. *et al. Neuron* **13**, 989-1002 (1994).
13. Morishima-Kawashima, M. *et al. J. Biol. Chem.* **270**, 823-829 (1995).
14. Goedert, M., Crowther, R. A. & Spillantini, M. G. *Neuron* **21**, 955-958 (1998).
15. Vincent, I., Jicha, G., Rosado, M. & Dickson, D. W. *J. Neurosci.* **17**, 3588-3598 (1997).
16. Nagy, Zs., Esiri, M. M., Cato, A.-M. & Smith, A. D. *Acta Neuropathol.* **94**, 6-15 (1997).
17. Busser, J., Geldmacher, D. S. & Herrup, K. *J. Neurosci.* **18**, 2801-2807 (1998).

erratum In Stefan Rahmstorf's News and Views article "Shifting seas in the greenhouse?" (*Nature* **399**, 523-524; 1999), the last five references were incorrectly numbered in the reference list. The correct numbering is:

10. Oppenheimer, M. *Nature* **393**, 325-332 (1998).
11. Boer, G. J., Flato, G. M. & Ramsden, D. *Clim. Dyn.* (submitted).
12. Hirst, A. C., Gordon, H. B. & O'Farrell, S. P. *Geophys. Res. Lett.* **23**, 3361-3364 (1996).
13. Knutson, T. R., Delworth, T. L., Dixon, K. & Stouffer, R. J. *J. Geophys. Res.* (submitted).
14. Hegerl, G. *et al. Clim. Dyn.* **13**, 631-634 (1997).

Daedalus

Metallic compost

All iron rusts in the end. Indeed, rapid rusting would be an ideal way of getting rid of old iron objects and structures. Rusting is a solid-state combustion; it absorbs oxygen and gives out heat — a reaction exploited in some heat-packs. So, says Daedalus, would an iron nail in an insulated oxygen flask get hotter and hotter as rusting started and speeded up, until it ultimately melted? Alas, no. Rusting is an electrolytic reaction, and needs liquid water. Once the temperature reaches 100 °C, rusting must stop.

But imagine a flask sealed and pressurized to 15 bar. Water inside it will remain liquid above 100 °C. At 15 bar it will reach 200 °C — at which a new rusting reaction starts to occur, one which works not only with water but with steam. Iron starts to reduce water directly, without oxygen, giving heat, rust and hydrogen. So an insulated scrap-iron dump, pressurized to 15 bar would, as it heated itself above 200 °C, and began to boil, would start to liberate this most ecologically virtuous of fuels. The temperature might even rise to the melting-point of iron.

So Daedalus's 'ferrous compost heap' is a large vertical cylinder, vacuum-insulated and with counter-current heat exchangers to feed the heat of its outputs back into its inputs. Iron scrap is fed into the top, and slowly descends into the increasingly hotter depths. Tappings at its base extract molten iron and iron oxide — the latter, of course, is high-grade iron ore. Air and steam are fed in at appropriate points, hydrogen and residual gases are extracted at others. It will be self-heating and self-maintaining.

At first Daedalus saw his ferrous compost heap as a simple disposer for old cans, car-bodies, oil drums and scrap metal too dirty, rusty or contaminated for normal recovery. But he now wonders what will happen to organic dirt and waste mixed in with the metal. A conventional incinerator oxidizes organic waste to steam, carbon dioxide and various pollutants. But iron is a strong reducing agent and (thanks to its variable valency) a versatile catalyst. It should reduce organic waste, perhaps to liquid or gaseous hydrocarbons or alcohols, useful fuels and solvents. Fed with a judicious mixture of old iron and general organic garbage, the ferrous compost heap should turn these nuisances into a range of handy fuels and raw materials, giving out heat but no pollution. Ecologists and industrialists alike would be delighted.

David Jones

Menstrual cycle alters face preference

Women prefer slightly feminized male facial shapes¹. Such faces (Fig. 1a) are given positive personality attributions¹ that might correlate with actual behaviour². In contrast, masculine features seem to signal immunological competence³. Heritable benefits can be realized only if conception follows copulation, so women might be more attentive to phenotypic markers indicating immunological competence during the follicular phase of the menstrual cycle when conception is most likely^{4,5}. Consistent with this hypothesis is the observation that women's preference for the odour of men with low fluctuating asymmetry (a correlate of testosterone-facilitated trait size and developmental stability) increases with the probability of conception across the menstrual cycle⁵. Symmetrical men report more extra-pair copulation partners⁶, and extra-pair copulation rates peak in mid-cycle⁷. Here we show that female preference for secondary sexual traits in male face shapes varies with the probability of conception across the menstrual cycle.

Japanese subjects reporting regular menstrual cycles and no use of oral contraceptives ($n=39$, mean age 21 years, mean cycle length 30 days) answered questions about the average cycle length and date of onset of previous menses. From this information, two testing sessions were arranged for the sampling of individuals' preferences during phases of both high and low conception risk in their menstrual cycles. Ovulation was assumed to occur 14 days before the onset of menses. Sessions after ovulation and before the onset of menses (the luteal phase) and sessions during menses were classified as 'low conception risk'. Sessions in the follicular phase (between the end of menses and ovulation) were classified as 'high conception risk'^{4,5,7}. Subjects were asked to select the face that they considered most 'physically attractive' ('miryoku-teki') from five Caucasian and, separately, five Japanese male faces (40% and 20% feminized, on average, and 20% and 40% masculinized). Subjects were also asked whether they currently had a 'steady boyfriend' ('tsukiatteiru'). Stimuli were symmetrical versions of those used (and illustrated) in previous research¹ but with hair visible, presented with full counterbalancing.

Repeated-measures analysis of variance (ANOVA) showed a significant main effect of conception risk (variance ratio ($F_{(1,37)}=9.47$; $P<0.004$), with subjects preferring faces that were less feminized in the high-conception-risk phase than in the low-conception-risk phase. No effects of stimuli origin (Caucasian or Japanese; Fig. 1b), partner or any interactions were found,

although there were trends indicating that women with a partner preferred faces that were more masculine ($F_{(1,37)}=3.59$, $P=0.066$) and underwent a greater cyclic change in preference than those without partners ($F_{(1,37)}=3.2$, $P=0.08$).

In a second experiment, an interactive method¹ was used that allowed British subjects (mean age 20 years) to alter composite male face shapes along continua ranging from 50% feminized to 50% masculinized. Subjects were asked to choose the most attractive face for a 'long-term relationship'

or a 'short-term sexual relationship'. Five facial continua (two from the experiment above¹ and three new ones; Fig. 1a) were prepared from separate symmetrical composites. Cycle and contraception details were obtained from 65 subjects, who completed three or four sessions at approximately weekly intervals: 27 made long-term and 28 made short-term judgements in all sessions. A further ten subjects made both judgements. Responses to faces were averaged over 'low-risk' and over 'high-risk' sessions.

Repeated-measures ANOVA for subjects not using oral contraception ('short-term', $n=17$; 'long-term', $n=20$; both judgements, $n=6$) showed no main effect of conception risk ($F_{(1,47)}=2.13$, $P=0.151$). However, conception risk interacted with type of relationship (short-term or long-term; $F_{(1,47)}=5.39$, $P=0.025$). For a short-term sexual relationship, the preferred face shape was less feminine during the high-conception-risk phase, whereas preferences remained constant when women judged attractiveness for a long-term relationship (Fig. 1c). Subjects preferred different levels of femininity in different composites ($F_{(4,188)}=11.1$, $P<0.01$), but the lack of significant interaction (risk $\times$ stimuli $\times$ relationship, $F_{(4,188)}=0.79$, $P=0.53$) showed that the cyclic change in preference was present for all stimuli. An analysis of data from subjects that were using oral contraception ($n=22$) revealed no cyclic changes in face shape preference ($F_{(4,96)}=0.001$, $P=0.97$) or interactions.

Dominance and quality as a parent are attributions made at opposite ends of the continuum relating to facial masculinity¹, and each might be associated with costs and benefits to reproductive success. A preference for males with a more masculine appearance might confer benefits for offspring in terms of resistance to disease but confer costs due to potentially decreased paternal investment.

In humans, concealed ovulation and limited visual similarity between offspring and their fathers⁸ can result in uncertainty of paternity. Such uncertainty, coupled with converging evidence for cyclic changes in female sexual behaviour and preferences for male characteristics^{5,7,9,10}, suggests that female mating strategy need not be entirely exclusive. As in some other species¹¹, selection might have favoured human females who pursued a mixed mating strategy under some ecological and social conditions. Women with a long-term sexual partner are more likely to have extra-pair copulations in the follicular phase of the cycle than during the luteal phase or menses⁷. A female might choose a primary

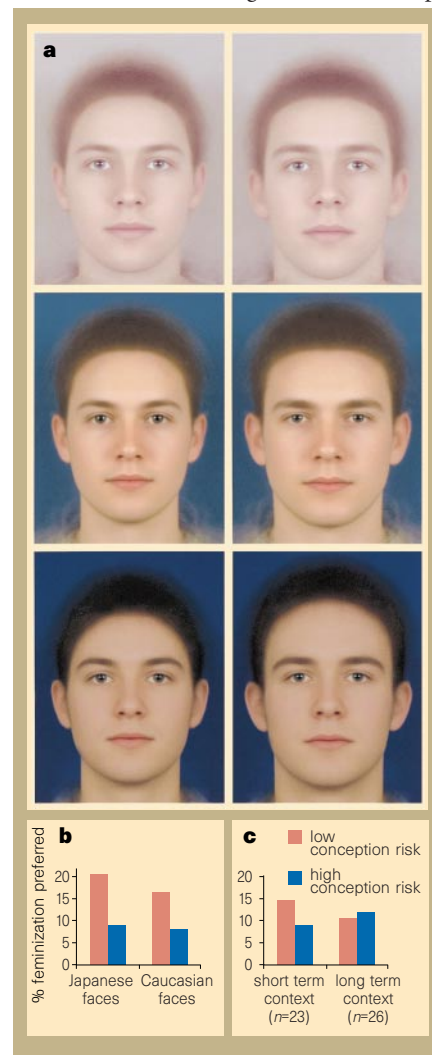


Figure 1 Cyclic shifts in the degree of femininity preferred in male faces. **a**, Face shapes that are 50% feminized (left) and 50% masculinized (right). Top: constructed¹ from 26 males, mean age 19.7 years; 37 females, 18.7 years. Centre: 21 males, 21.0 years; 40 females, 21.0 years. Bottom: 18 males, 19.8 years; 38 females, 20.8 years. **b**, Mean feminization preferred in Japanese and Caucasian composites by Japanese subjects ($n=39$) in high- and low-conception-risk phases. **c**, Mean feminization preferred across faces for short- and long-term conditions (experiment 2) in high-risk and low-risk phases.

partner whose low masculine appearance suggests cooperation in parental care ('long-term' preferences are unchanged across the menstrual cycle) but occasionally copulate with a male with a more masculine appearance (indicating good immunocompetence) when conception is most likely. Sexual behaviour arising from cyclic preferences might allow individuals to accrue benefits from polyandry while maintaining the advantage of ostensive monandry.

I. S. Penton-Voak*, D. I. Perrett*, D. L. Castles†‡, T. Kobayashi†, D. M. Burt*, L. K. Murray*, R. Minamisawa†

*School of Psychology, University of St Andrews, Fife KY16 9JU, UK

e-mail: ip@st-and.ac.uk

†Hasegawa Laboratory,

Department of Life Sciences, University of Tokyo, Komaba, Tokyo 153, Japan

‡School of Life Sciences,

Roehampton Institute London, Whitelands College, London SW15 3SN, UK

1. Perrett, D. I. *et al.* *Nature* **394**, 884–887 (1998).
2. Berry, D. S. & Wero, J. L. F. *J. Pers.* **61**, 497–523 (1993).
3. Folstad, I. & Karter, A. *Am. Nat.* **139**, 603–622 (1992).
4. Regan, P. C. *Can. J. Hum. Sexual.* **5**, 145–156 (1996).
5. Gangestad, S. W. & Thornhill, R. *Proc. R. Soc. Lond. B* **265**, 927–933 (1998).
6. Gangestad, S. W. & Thornhill, R. *Evol. Hum. Biol.* **18**, 69–88 (1997).
7. Baker, R. R. & Bellis, M. A. *Human Sperm Competition: Copulation, Masturbation and Infidelity* (Chapman & Hall, London, 1995).
8. Christenfeld, N. & Hill, E. A. *Nature* **378**, 669 (1995).
9. Frost, P. *Percept. Motor Skills* **79**, 507–514 (1994).
10. Grammer, K. *Ethol. Sociobiol.* **14**, 201–208 (1993).
11. Graves, J., Ortegaruano, J. & Slater, P. J. B. *Proc. R. Soc. Lond. B* **253**, 3–7 (1993).

Limbs move beyond the Radical fringe

The *fringe* genes and the Notch signalling pathway are important in limb development in vertebrates and in *Drosophila*, helping to establish the 'organizing centres' that are required for the proximal–distal outgrowth of appendages^{1–5}. Three vertebrate *Fringe* genes have been identified: *Manic* (*Mfng*), *Lunatic* (*Lfng*) and *Radical fringe* (*Rfng*)^{4–8}. Here we show that the mouse *Rfng* gene is not required for limb development, even though it is expressed in the developing limb bud. But we have found that several developmental defects, which at first we attributed to a loss of *Rfng* function after mutating the gene, in fact arose as a result of the insertion of a selection cassette into the *Rfng* locus that affects the expression of a neighbouring gene or genes. Our findings highlight the need for optimization in designing mutant alleles to probe developmental processes^{9,10}.

Studies of *Rfng* expression in chick limb mutants and gain-of-function experiments in limb buds using retroviral misexpression

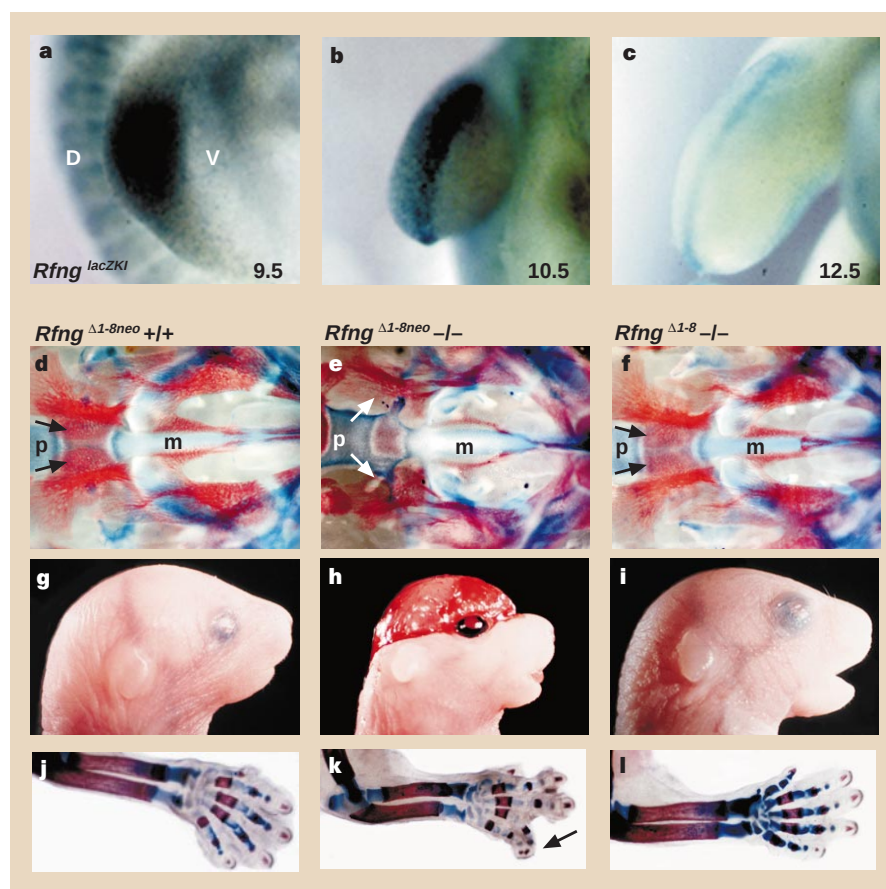


Figure 1 *Rfng* is expressed in the mouse limb bud and is not required for limb development. *Rfng* expression in *Rfng*^{lacZKI} heterozygote limbs, as shown by X-gal staining, at: **a**, 9.5 d.p.c.; **b**, 10.5 d.p.c.; and **c**, 12.5 d.p.c. Dorsal (D), left; ventral (V), right. Replacement of the *Rfng* locus with a PGKneo-selection cassette results in pleiotropic developmental defects. **d**, **g**, **j**, Wild-type fetuses; **e**, **h**, **k**, *Rfng*^{Δ1-8neo} homozygotes; and **f**, **i**, **l**, *Rfng*^{Δ1-8} homozygotes at 18.5 d.p.c. **d–f**, Skeletal staining of the secondary palate. Anterior is to the right. **d**, The palatal (p; indicated by arrows) and maxillary (m) shelves are nearly fused. **e**, *Rfng*^{Δ1-8neo} homozygote palatal shelves (p; white arrows) have not elevated towards the midline, and the maxillary shelves (m) are wider than in the wild type. **f**, *Rfng*^{Δ1-8} homozygote, showing normal palatal (p) and maxillary (m) shelf positioning. **g**, Wild-type fetus with closed cranium. **h**, *Rfng*^{Δ1-8neo} homozygote, showing exencephaly. **i**, *Rfng*^{Δ1-8} homozygotes are indistinguishable from wild type (**g**). **j–l**, Skeletal staining of the forelimbs. Anterior is up, posterior is down. **j**, *Rfng*^{Δ1-8neo} wild-type limb has five digits. **k**, *Rfng*^{Δ1-8neo} homozygote, showing soft-tissue syndactyly of digits 4 and 5 (arrow). **l**, *Rfng*^{Δ1-8} homozygotes are indistinguishable from wild type (**j**).

both indicated that the juxtaposition of cells that did and did not express *Rfng* is sufficient for the formation of the apical ectodermal ridge (AER), an organizing centre that forms at the boundary between dorsal and ventral cells at the distal edge of the limb bud^{4,5}. We have used homologous recombination in mouse embryonic stem cells to investigate *Rfng* expression in the mouse limb bud and to test whether it is necessary for the formation of the AER.

We inserted the *Escherichia coli lacZ* gene into the *Rfng* genomic locus to create the *Rfng*^{lacZKI} 'knock-in' allele. Analysis of β-galactosidase activity in *Rfng*^{lacZKI} limb buds revealed expression in the developing and mature AER (Fig. 1a–c). To create a null allele, *Rfng*^{Δ1-8neo}, we deleted 2.4 kilobases of the *Rfng* genomic locus containing the entire coding region (exons 1–8) and replaced it with a 'PGKneo-selection' cassette (which confers a resistance to the

antibiotic neomycin and acts as a means for selecting these clones) in the opposite transcriptional orientation.

No homozygous *Rfng*^{Δ1-8neo}-deficient mouse survived past birth ($n = 244$ progeny from heterozygote intercrosses on a mixed 129P2 × C57BL/6J hybrid background; $n = 70$ progeny from heterozygote intercrosses on a mixed 129P2 × 129S6 hybrid background). The mutant mice displayed several developmental defects, including a cleft palate and reduced mandible (64%), exencephaly (24%), and forelimb defects (12%) consisting of soft-tissue syndactyly or oligodactyly (total $n = 225$ on a mixed 129P2/OlaHsd × C57BL/6J hybrid background, examined between 18 days post-coitum (d.p.c.) and birth; Fig. 1).

We went on to remove the PGKneo-selection cassette to create the *Rfng*^{Δ1-8} null allele by two independent methods: mating *Rfng*^{Δ1-8neo} mice to mice that expressed Cre

recombinase (FVB/N genetic background)¹¹ and microinjection of a circularized Cre recombinase-expressing plasmid into *Rfng*^{Δ1-8neo} heterozygote fertilized eggs (129P2/OlaHsd × C57BL/6J hybrid background). In contrast to *Rfng*^{Δ1-8neo} homozygotes, *Rfng*^{Δ1-8} homozygotes on both genetic backgrounds were viable and were morphologically wild type (total *n* = 172; Fig. 1f, i, l). We conclude that *Rfng* function is not essential for AER formation or for proximal-distal limb outgrowth¹². The absence of a limb phenotype in the *Rfng* null allele may be due to functional overlap between Fringe family members.

We suggest that the phenotypes associated with the introduction of a PGKneo-selection cassette into the *Rfng* locus (*Rfng*^{Δ1-8neo} allele) are caused by interference with the proper expression of a neighbouring gene or genes whose expression overlaps with that of *Rfng* and are required for neural tube, craniofacial and limb development. There have been previous examples in which PGKneo effects have been examined. In such cases, the target genes have been part of a related gene complex that shares both regulatory control and functions (for example, the myogenic regulatory factors, HoxD and globin gene clusters¹³).

Our findings with *Rfng*-targeted alleles may be relevant for interpreting results from two *Lfng*-targeted alleles, *Lfng*^{lacZ} and *Lfng*^{Δ1neo} (refs 9,10). In both of these *Lfng* gene-targeted alleles, the PGKneo-selection cassette was not excised from the locus^{9,10}. The *Lfng*-targeted mutants exhibited phenotypic defects restricted to somitogenesis; however, there were unexplained differences in somite epithelialization. This phenotypic difference was attributed to either differences in genetic background or to differences in the design of the targeting constructs, although both targeted alleles were thought to be null alleles^{9,10}.

With these *Lfng* alleles, it might therefore be possible that the PGKneo cassette affects the expression of a neighbouring gene(s) (perhaps paralogous between *fringe* family members). One such candidate is *Uncx4.1*, a paired type homeodomain gene that is physically linked to *Lfng* and whose expression in the caudal half of condensing somites is altered in the *Lfng* mutants^{9,10,14}. Our findings reinforce the importance of understanding the increasingly apparent presence of embedded and shared regulatory elements in genomes¹¹.

Jennifer L. Moran, John M. Levorse,
Thomas F. Vogt

Department of Molecular Biology,
Lewis Thomas Laboratory, Princeton University,
Princeton, New Jersey 08544, USA
e-mail: tvogt@molbio.princeton.edu

3. Panin, V. M., Papayannopoulos, V., Wilson, R. & Irvine, K. D. *Nature* **387**, 908–912 (1997).
4. Laufer, E. et al. *Nature* **386**, 366–373 (1997).
5. Rodriguez-Esteban, C. et al. *Nature* **386**, 360–366 (1997).
6. Wu, J. Y., Wen, L., Zhang, W.-J. & Rao, Y. *Science* **273**, 355–358 (1996).
7. Johnston, S. H. et al. *Development* **124**, 2245–2254 (1997).
8. Cohen, B. et al. *Nature Genet.* **16**, 283–288 (1997).
9. Zhang, N. & Gridley, T. *Nature* **394**, 374–377 (1998).
10. Evrard, Y. A. et al. *Nature* **394**, 377–381 (1998).
11. Lakso, M. et al. *Proc. Natl Acad. Sci. USA* **93**, 5860–5865 (1996).
12. Christen, B. & Slack, J. M. W. *Nature* **395**, 230–231 (1998).
13. Olson, E. N., Arnold, H. H., Rigby, P. W. J. & Wold, B. J. *Cell* **85**, 1–4 (1996).
14. Moran, J. L. et al. *Mamm. Genome* **10**, 535–541 (1999).

Zhang and Gridley reply — We too have created mice that have a large deletion in the *Radical fringe* gene and investigated the resulting effect on their development. Our results agree with those of Moran et al. in showing that *Rfng* is not an essential gene in mice, but they do not support the idea that the presence of a *neo*-selection cassette inserted into the *Rfng* locus invariably leads to embryonic lethality.

We constructed and analysed a targeted mutation of the *Rfng* gene by using a targeting vector in which a section of the *Rfng* genomic sequence, encoding the region from amino acid 52 of the *Rfng* protein through to the end of the *Rfng* gene, was deleted. The vector contained a *neo*-selection cassette in the same transcriptional orientation as the *Rfng* gene. Mice homozygous for our *Rfng* mutant allele formed viable, morphologically normal adults at the expected mendelian frequencies (34 of 126 homozygotes (27%) on a mixed C57BL/6J × 129/SvImJ background, and 3 of 16 homozygotes (19%) on an inbred 129/SvImJ background). The absence of wild-type *Rfng* transcripts in our *Rfng* homozygous-mutant adult mice was confirmed by northern blotting.

It is possible that a difference in the construction of the two targeted *Rfng* alleles could explain the different mouse phenotypes generated by Moran et al. — the orientations of the *neo*-selection cassette were not the same in the two targeted *Rfng* alleles, for instance. Reversing the orientation of the *neo* sequence for transcription creates a potent splice-acceptor site¹².

In considering the implications of their finding for the interpretation of *Lfng* mutant phenotypes^{3,4}, Moran et al. correctly point out that, in some instances, the presence of a *neo*-selection cassette in a targeted allele can affect the transcription of other genes, and hence the mutant phenotypes they generate. To confirm that the mutant phenotype is independent of the presence of a *neo* cassette in the two *Lfng* mutant alleles, a targeted *Lfng* allele constructed without a *neo* cassette needs to be tested.

However, we believe that it is unlikely that the *neo* cassette in the two *Lfng* mutant alleles is responsible for the alterations in *Uncx4.1* transcription or for the defects in

somite formation that we detected in *Lfng* homozygous-mutant embryos. The spontaneous mouse mutant *pudgy* (*pu*) is caused by a frameshift mutation in the *Dll3* gene⁵, which encodes a ligand for the Notch family of proteins. This frameshift leads to premature truncation of the *Dll3* protein⁵. As homozygous *Dll3*^{pu}/*Dll3*^{pu} mutant embryos show defects in somite formation very like those seen in *Lfng* mutant embryos, we compared *Uncx4.1* transcription in both *Dll3*^{pu}/*Dll3*^{pu} and *Lfng*^{lacZ}/*Lfng*^{lacZ} mutant embryos and found that the *Uncx4.1* expression pattern was similarly disorganized in both mutants. Given that the presence of a *neo* cassette does not alter *Uncx4.1* expression in *Dll3*^{pu}/*Dll3*^{pu} embryos, the simplest explanation is that the alterations in *Uncx4.1* expression and somite formation in *Lfng*^{lacZ}/*Lfng*^{lacZ} mutant embryos are caused by perturbations of the Notch signalling pathway, and not by the presence of a *neo* cassette in the *Lfng*^{lacZ} mutant allele.

Nian Zhang, Thomas Gridley

The Jackson Laboratory,
Bar Harbor, Maine 04609-1500, USA
e-mail: gridley@jax.org

1. Jacks, T. et al. *Nature Genet.* **7**, 353–361 (1994).
2. Carmeliet, P. et al. *Nature* **380**, 435–439 (1996).
3. Zhang, N. & Gridley, T. *Nature* **394**, 374–377 (1998).
4. Evrard, Y. A. et al. *Nature* **394**, 377–381 (1998).
5. Kusumi, K. et al. *Nature Genet.* **19**, 274–278 (1998).

Risk of collisions for constellation satellites

The increasing quantity of debris orbiting the Earth is causing concern for space agencies. At the average collision velocity of 10 km s⁻¹, even projectiles 1 cm across can damage satellites. Impacts have a significant chance of occurring, especially for large structures that remain in orbit for a long time¹. Particularly at risk are satellites at altitudes of about 800 and 1,400 km, where there is already a high density of orbiting bodies and atmospheric drag is not effective in removing small fragments. We find that the dynamical architecture of satellite constellations, several of which are to be launched over the next decade, makes them particularly vulnerable to the consequences of an impact, which may set up a kind of chain reaction, triggering more collisions.

There are currently about 10⁴ objects larger than 10–20 cm across orbiting the Earth, with 10⁵ objects more than 1 cm across, and 10⁷ bodies exceeding 1 mm in diameter. These have a total mass of about 3 × 10⁶ kg and a cross-section of 4 × 10⁴ m² (refs 2,3), and the average flux (the expected impact rate per projectile and per unit target cross-section) is about 3 × 10⁻¹⁰ m⁻² yr⁻¹. In the long term, debris from collisions with

1. Shubin, N., Tabin, C. & Carroll, S. *Nature* **388**, 639–648 (1997).
2. Irvine, K. D. & Vogt, T. F. *Curr. Opin. Cell Biol.* **9**, 867–876 (1997).

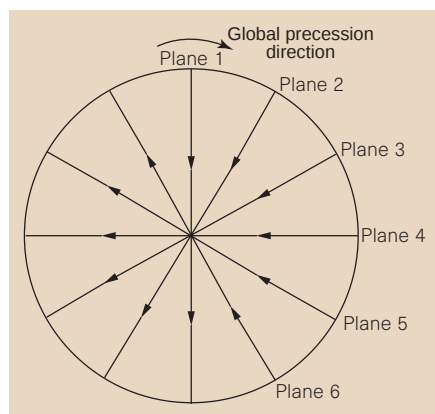


Figure 1 The orientation of the six orbital planes of an Iridium-like constellation, as seen from the celestial north pole. The revolution of the satellites is indicated by the arrows. Note the 'counter-rotating' planes 1 and 6. If a satellite orbiting in either of these planes is disrupted, its fragments will pose an impact hazard much higher than average to the satellites in the other plane.

these objects may irreversibly pollute portions of near-Earth space^{4,5}.

Satellite constellations, consisting of tens to hundreds of satellites orbiting at about the same altitude, will be launched into orbit over the next ten years, mainly for global telecommunication purposes⁶. One of these, the Iridium system, is already operational. It consists of 66 satellites (plus six spares) orbiting in six different orbital planes at an altitude of about 780 km and with an inclination of 86.4 degrees. The wet mass of each satellite is 667 kg. Although some measures have been taken to minimize the risk of impact, the possibility of disruptive collisions cannot be ruled out, as the selected altitude corresponds to a peak of the expected debris flux. Assuming a catastrophic threshold for the impact specific energy of $4.5 \times 10^3 \text{ J kg}^{-1}$, as inferred from laboratory experiments on spacecraft mock-ups⁷, the flux of potentially shattering fragments at 800 km is about $10^{-5} \text{ m}^{-2} \text{ yr}^{-1}$. For a constellation cross-section of about 10^3 m^2 , the risk of a catastrophic impact is about 10% per decade.

The architecture of such a constellation will make a break-up event particularly dangerous, owing to the spreading of the resulting fragment swarms, on a timescale intermediate between those analysed in previous studies (several hours⁸ and several decades⁹). This is particularly true when differential precession of the orbits leads the fragments to encounter satellites revolving around the Earth in the opposite direction (Fig. 1); this makes head-on collisions possible, with higher impact speeds and greater collision probability.

Results of a simulation of an Iridium satellite being broken up by a 1-kg projectile at a relative speed of 10 km s^{-1} are shown in Fig. 2. We have modelled the resulting

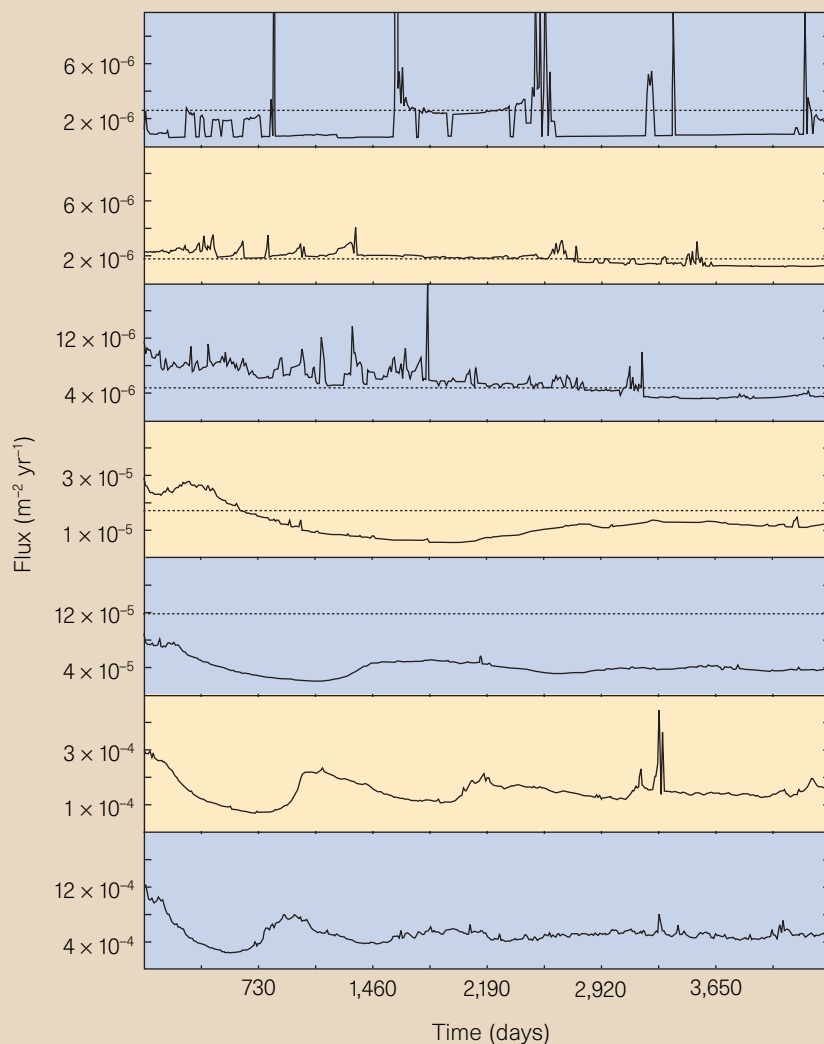


Figure 2 Flux of fragments over time produced by the simulated impact break-up of an Iridium satellite orbiting in plane number 6 of Fig. 1. Panels show, from top to bottom, the fluxes at decreasing kinetic energy, E : $2E > 10^9 \text{ J}$; $10^8 < 2E < 10^9 \text{ J}$; $10^7 < 2E < 10^8 \text{ J}$; $10^6 < 2E < 10^7 \text{ J}$; $10^5 < 2E < 10^6 \text{ J}$; $10^4 < 2E < 10^5 \text{ J}$; and $10^3 < 2E < 10^4 \text{ J}$. The target satellite is located in plane number 1 (Fig. 1). Horizontal dotted lines represent background flux in the same energy ranges, calculated from the overall space debris population, according to ref. 2.

fragment swarm (8,000 bodies heavier than 1 g) under Earth's oblateness and air drag perturbations, and computed as a function of time the cumulative probability of impact with another constellation satellite moving in the opposite direction.

For an impact energy between 10^7 and 10^8 J , corresponding to such projectiles, the collision probability stays higher than the background level from the general orbiting population for several years. This means that, after the initial break-up, the probability that a second one will follow within five years is about 10%. This may then trigger a chain-reaction effect with a characteristic timescale of about a century, much less than the current estimates with the general debris population (300 to 500 yr)¹⁰.

A. Rossi*, G. B. Valsecchi†, P. Farinella‡

*CNUCE-CNR, Via S. Maria 36,
56126 Pisa, Italy

e-mail: a.rossi@cnuce.cnr.it

†IAS-CNR, Via Fosso del Cavaliere,
00133 Roma, Italy

‡Dipartimento di Astronomia,
Universita' di Trieste, 34131 Trieste, Italy

- Rossi, A. & Farinella, P. *ESA J.* **16**, 339–348 (1992).
- Pardini, C., Anselmo, L., Rossi, A., Cordelli, A. & Farinella, P. *Adv. Astronaut. Sci.* **99**, 1041–1058 (1998).
- Kessler, D. J. *et al.* NASA Technical Memorandum 104825 (NASA, 1996).
- Kessler, D. J. & Cour-Palais, B. G. *J. Geophys. Res.* **83**, 2637–2646 (1978).
- Rossi, A., Cordelli, A., Farinella, P. & Anselmo, L. *J. Geophys. Res.* **99**, 23195–23210 (1994).
- van der Ha, J. C. (ed.) *Mission Design and Implementation of Satellite Constellations* (Kluwer, Dordrecht, 1998).
- McKnight, D. S., Maher, R. L. & Nagl, L. *Fragmentation Algorithms for Satellite Targets (FAST) Empirical Breakup Model, Version 2* (Kaman Sciences, 1992).
- Walker, R., Crowther, R., Wilkinson, J., Stokes, P. H. & Swinerd, G. G. in *Mission Design and Implementation of Satellite Constellations* (ed. van der Ha, J. C.) 317–326 (Kluwer, Dordrecht, 1998).
- Rossi, A., Pardini, C., Anselmo, L., Cordelli, A. & Farinella, P. in *Mission Design and Implementation of Satellite Constellations* (ed. van der Ha, J. C.) 327–335 (Kluwer, Dordrecht, 1998).
- Cordelli, A., Farinella, P. & Rossi, A. *Planet. Space Sci.* **46**, 691–699 (1998).

When the means do not justify the end

Sudden Origins: Fossils, Genes, and the Emergence of Species

by Jeffrey H. Schwartz
Wiley: 1999. 408 pp. £19.99, \$27.95

Eörs Szathmáry

The origin of species has long fascinated biologists. Although Darwin's major work bears it as a title, it does not provide a solution to the problem. Does Jeffrey Schwartz give one? I am afraid that, in general, he does not, but the book is still interesting.

Schwartz presents a detailed and informative historical account of evolutionary biology. In fact, the book could be read as a history of evolution, and will probably occupy such a niche. Schwartz is much more ambitious than this, however: he wants to convince the reader that his "new evolution" is the Holy Grail of the field.

These are the elements of his proposal. First, forces of speciation are largely independent of selection within the species. Second, speciation is typically triggered by a mutation in a homeobox gene. Third, such a mutation is likely to become fixed in population isolates. Fourth, the mutation remains unnoticed for a long time because of its recessive nature. Fifth, when it reaches homozygosity in the small isolate, it will appear in several individuals simultaneously. Sixth, the mutation, since it affects development, is likely to affect mate choice as well; hence, mutant homozygotes are likely to mate among themselves.

Although Schwartz has in one sense resurrected Richard Goldschmidt's "hopeful monsters" (organisms with radically altered characteristics, produced in single large steps), he thinks that he has managed to avoid the associated pitfalls. The homeobox mutations would be macromutations in their effects, but from the genetic point of view they would qualify as micromutations, since no large genomic rearrangements are implied. The problems of finding a mate also seem to be solved by the dynamics of recessives (mutant homozygotes are thought to appear simultaneously, following a lag phase of heterozygosity). And, since homozygosity and preferential mating (pre-zygotic isolation) go hand in hand, there would be no transitional forms.

A drawback of this scenario is the absence of a population-genetic treatment. One is therefore left uncertain about the effects of population size, fixation time, and so on. It is not even certain whether Schwartz thinks that the mutant homozygotes have a selective advantage or are effectively neutral for the individual. In the latter case, the logic of the scenario still suggests that they would influence the dynamics of sexual selection,



Evolution by radical mutation: "Tragic anatomies" by Jake and Dinos Chapman.

an interesting possibility. There is nothing wrong with dreaming up a verbal evolutionary scenario, but these days this surely cannot be the last word.

Another shortcoming of the book is the lack of at least one documented case of speciation that may have happened following Schwartz's scenario. Examples that could be taken as evidence for the component processes are discussed in some detail, but this does not rule out the possibility that one case that fits the complete scenario would never be found.

Schwartz shares a distrust of selection with some contemporary biologists. This is, I think, why he chose to ignore, for example, Peter Sheldon's work on the gradualism apparent in the evolution of the pygidial ribs of trilobites. When he discusses the celebrated computer study by Dan Nilsson and

Susanne Pelger on the evolution of the optical structure of a fish eye, he mistakenly suggests that the intermediates are not selective improvements on the previous forms. It is revealing that he dismisses this scenario by saying: "Do we actually need to invoke such an elaborate thought experiment in order to understand the origin of the vertebrate eye, or any eye, for that matter? I think not. And the reasons lie in knowing that there are homeobox genes for eye formation and that when one of them, the *Rx* gene in particular, is activated in the right place and at the right time, an individual has an eye."

This is utterly misleading. Schwartz ignores the fact that homeobox genes are selector genes. They can do nothing if the genes regulated by them are not there. It is these genes that specify in detail the adaptive structure of the organs. To be sure, turning on a homeobox gene at the wrong place can result in the appearance of an ectopic organ, but only if the genes for that organ are present in the same individual. It is totally wrong to imply that an eye could be produced by a macromutation when no eye was ever present in the lineage before. Homeotic mutations that reshuffle parts do happen, and sometimes they may have led to fixation of real evolutionary novelties, but this does not mean that such changes are implied in the majority of speciations. In fact, macromutations of this sort are probably frequently maladaptive, in contrast to the vast number of past and present species — not to mention the fact that morphological differences between related species can be minute.

The history given in the book is fascinating, and some of the suggestions merit further work, but the reader is likely to be deeply disappointed more than once before the end. □

Eörs Szathmáry is at the Collegium Budapest, Szentháromság u2, H-1014 Budapest, Hungary.

This passion of our kind

La Galassia Mente

by Rita Levi Montalcini
Baldini & Castoldi: 1999. 218 pp. Lire 29,000
Lucia Galli-Resta

The Galaxy of the Mind aims to stimulate reflection on the complexity of the human brain. At a time when biotechnology has such a powerful impact on our lives and expectations, in a country where scientific education has still to be 'perfected', this book has its small, valuable place. The book has a simple and important message: our mind,

about which we still know so little, is our master and our slave. The values we foster in our culture and in our children will ultimately decide whether the mind serves knowledge or intolerance, and how it will guide mankind in the use of knowledge.

Rita Levi Montalcini begins by tracing a succinct picture of the evolutionary origins of the nervous system, delineating its principal unitary characteristics as well as its diversification. Her description culminates with Lucy: "52 bones ... of the skeleton of a single individual, a female hominid ... with a skull not larger than a coconut". Here, halfway through the book, I switched off the light and went to sleep. In the dark of the night I wondered whether angst moved in her brain and whether she, as W. H. Auden once wrote, "our own shrew ancestor, was Nobody, but still could take herself for granted".

Resuming the next day, I met the 'human brain' in the lively metaphor of a game of chess — neocortex as king; limbic system as queen; and the subcortical regions as their bishops, castles and knights. Last but not least, the pawns: lymphokines, growth factors, enkephalins, and so on. In the most difficult chapters of the book, language, thought, mind and consciousness are all briefly considered in a summary of views which one might or might not share, but which still convey the vividness of the debate, and how far we are from a solution.

The author has the merit of rarely taking sides, doing so mostly through the verses chosen to introduce each chapter. Although reading these chapters may be somewhat laborious, the book rushes on. On the subjects of free will and emotions, the author's charm mitigates the reader's frustration at the brief treatment of the subjects, and leaves us with two images. The first is a dog roaming at the end of a long leash, to suggest the interplay — limited yet vital — between what we call our freedom of thought and our biological constraints. The second image is a little warning, showing the folded mantle of the neocortex wrapping, but not trapping, a primordial reptilian brain, as powerful in giving us emotions as it is fearsome in its obscure potentialities.

The finale, which is prefigured by a word appearing often in the text — ethic — is a simple message of great relevance: science has revealed new and unpredictable horizons to mankind; the choice is between good and evil. As Levi Montalcini says: "The bond between science and morality must be strengthened."

To return to Auden, "this passion of our kind/ for the process of finding out/ is a fact one can hardly doubt". Levi Montalcini's book suggests a healthy diet to feed this passion: "science allows an unambiguous language, understandable by the most different nations ... Promoting a friendship based on the exchange of knowledge and reciprocal

contacts may be the most important contribution of science and technology to humankind."

One may like or dislike this little book aimed at a popular audience, but the message — from a great scientist, and a woman who is a living embodiment of the passion for knowledge — should not be ignored. □

Lucia Galli-Resta is at the Istituto di Neurofisiologia del CNR, via San Zeno 51, 56127 Pisa, Italy.

Ditties of the fragile and the inscrutable

Musical Networks: Parallel Distributed Perception and Performance

edited by Niall Griffith and Peter M. Todd
MIT Press: 1999. 385 pp. \$37.50, £26.50

Christopher Longuet-Higgins

From time to time, a previously dormant subject suddenly comes to life and challenges our preconceptions about its concepts and methods. A quarter of a century ago this happened to music theory, when people realized that the perception and performance of music are psychological processes that might be amenable to computational modelling.

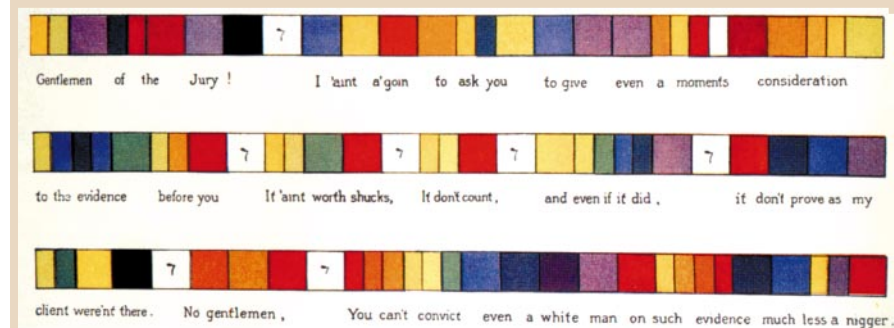
A good start was made in applying the concepts of artificial intelligence to music psychology, but this was soon eclipsed by the connectionist paradigm — the doctrine that, since the brain is a parallel distributed processor or, as some would say, an adaptive neural network, all statements about music should be reducible to statements about the activity of such networks.

Sidney Brenner, one of the founders of molecular biology, once remarked that a mature science was one in which the number of competing theories was equal to one. On that criterion, music psychology is still in its

infancy. There is fierce competition, not only between theories, but between types of theory — between fragile artificial-intelligence models and inscrutable neural networks. The AI models seem more successful at capturing the role of knowledge in musical cognition, the neural-net models at explaining the acquisition of perceptual skills.

Musical Networks testifies to the breadth of influence that connectionism has exerted on the psychology of music, or at least on its literature. Many of the chapters were originally published in a symposium entitled *Music and Creativity* (edited by Griffith and Todd and published in a special issue of *Connection Science*, Vol. 6, parts 2 & 3; 1994), but here we find connectionist accounts of processes as wide-ranging as pitch perception, key identification, rhythmic entrainment, voice separation, metre induction and even the use of Boltzmann machines in the harmonization of chorales. One is reminded of the creativity test, "How many uses can you think of for a brick?" Unfortunately, the non-specialist is left in no position to judge which of the many networks proposed are realized in the central nervous system and which are not.

A topic conspicuous by its absence from *Musical Networks* is musical grammar. There are two likely reasons for this. One is a suspicion that the topic might have been pre-empted (actually it wasn't) by Fred Lerdahl and Ray Jackendoff's *A Generative Theory of Tonal Music* (MIT Press), an attempt to do for music what Noam Chomsky did for language. The other is the connectionist's allergy to replacement rules and other top-down procedures for arriving at well-formed strings, sequences, sentences and phrases. The great attraction of connectionism is that it seems to solve the problem of teaching a system new tricks: let it pick them up for itself. The snag is, of course, that the learning process may be impossible to understand. For practical purposes, such as music



A sight of sound

"Lawyer Spoke Stith's address to the jury" (1900) might be a colourful representation of synaesthesia. Shown above is a visualization of the colours of his spoken words by the American

architect and designer E. J. Lind, who formulated such ideas in the 1880s. Taken from *Colour and Meaning: Art, Science and Symbolism* by John Gage (Thames & Hudson, £32).

technology, incomprehensibility may not matter. But to the theoretical psychologist it is a challenge that has already led to the introduction of new methods of data analysis, and one may look forward to an exciting moment when top-down and bottom-up approaches to the musical mind finally meet in the middle. □

Christopher Longuet-Higgins is at the Centre for Research in Perception & Cognition, University of Sussex, Falmer, Brighton BN1 9QY, UK.

The world as a patchwork

Metapopulation Ecology

by Ilkka Hanski

Oxford University Press: 1999. 313 pp.
£22.50 (pbk)

Camille Parmesan

As we field biologists see our favourite study sites succumb to anthropogenic assaults, we must increasingly spend our time searching for 'pristine' habitats in which to study natural ecological phenomena. They are nearly impossible to find, but, if found, are cherished for whatever information can be gleaned before they, too, become victims of the plough or the axe, acid rain or nitrogen run-off. Of necessity, ecologists have either partially or wholly redirected their research towards the effects of these and other anthropogenic disturbances on remaining 'natural' systems.

This shift has not come easily to those trained in classical ecology — the purist's sentiment that basic research is superior to these more 'applied' questions is alive in all ecological sub-disciplines. Fortunately, many researchers (at all levels) now regard such classical idealism as snobbery at best, and harmfully delusional at worst. Ilkka Hanski shows his colours by stressing practical applications throughout this *tour de force* exposition of the metapopulation approach to ecology, evolution and conservation biology.

The metapopulation approach recognizes the influence that otherwise discrete populations can have on one another when neighbouring patches exchange small numbers of migrants. This network of populations, when viewed in its own right, may express very different dynamics from those of its separate population components.

The accelerating growth of this field, fuelled in large part by Hanski himself, has ecologists in other areas looking nervously askance. How much bigger can this bouncing baby become? Will it come to define conservation biology? Can the world be finally understood only if we consent to view it through an Esheresque maze of ephemeral patches set within networks of even more patches?

Fear not. Here lies neither a diatribe nor a plea. Hanski meticulously details the history,

growth and empirical and theoretical underpinnings of metapopulation biology. He approaches the subject in a down-to-earth manner, concentrating on his prime motivation — to present a variety of concepts and models which may help us to understand and conserve what still remains. Being equally astute in theory and empiricism, he manages this with aplomb.

While admitting that the metapopulation approach will not apply in every instance, Hanski attacks its sceptics in a lively manner: "I want to make the point that ignoring the actual spatial population structure of any species, at whatever the scale that structure might occur, would be an exceedingly silly thing to do." Bravo! Leave the philosophically inclined to debate the delineation of a population, and let's get on with the business at hand.

And serious business it is — this book is not for the faint-hearted. The density of information, page after page, is rather like a Yorkshireman's description of a good cup of tea — thick enough to stand a spoon in. Hanski's attempt at making his work accessible to the bulk of population biologists is generally successful, and impressively eloquent in arguments he considers crucial. However, he also assumes a high degree of knowledge, as well as motivation to understand the subject. I would suggest new graduate students have a basic ecology text or dictionary handy to follow the arguments fully. The casual reader can still come away somewhat enlightened, as plainly written summaries begin and end each chapter, and often each section. Frankly, I found his treatment a refreshing refusal to cater to the lowest common denominator.

Unfortunately, Hanski's impressive style is ill-supported by abysmal print quality. The blurred outlines of the letters complemented the small, overly-serifed font to give me a solid headache within 20 minutes of reading. I thought I needed glasses, when in fact I needed a different production editor.

I applaud Hanski for embracing what many ecologists emotionally reject: the incorporation of evolution into what has largely been an ecological domain. Indeed, the few empirical genetic studies that do exist bolster his assertion that "local adaptation influences spatial dynamics and spatial dynamics influence local dynamics". However, evolutionists must be prepared for an unconventional use of "fitness" to refer to population growth rate (as well as to individual fitness) — a usage justified by model results which show that population growth rate can be important to evolution at the metapopulation scale.

I learned a lot in areas I knew dimly and was entertained in areas I knew well — a sure sign of a good book. Hanski's emphasis on the application of models, particularly to conservation biology, will make this work not only a classic for academic population biologists,

but an essential reference for managers and conservation planners as well. □

Camille Parmesan is at Integrative Biology, Patterson Laboratories, University of Texas, Austin, Texas 78712, USA.

Taking a gender tiger by the tail

Why So Slow? The Advancement of Women

by Virginia Valian

MIT Press: 1999. 424 pp. \$17.50, £12.50 (pbk)

Christine Wennerås and Agnes Wold

Despite the large influx of women to universities during the past three decades, women's academic careers are generally far less brilliant than those of their male peers. Common prejudice tells us that this is largely due to 'women's special priorities', for example that women nurture their families rather than their careers. In fact, married women with children are at least as scientifically productive as single women, and advance equally, or more, rapidly than their childless female colleagues. Further, no simple relation between the role of women in society at large and their capacity to climb the academic career ladder can be found. For example, the 'women friendly' Scandinavian countries have excellent child care and a high proportion of women politicians, but a glaring scarcity of female professors.

The reasons for the slow advancement of women in the academic system must be sought within the system itself. Resources necessary for doing research — positions, laboratory space, personnel and money — are limited, and their distribution is as a rule based on the judgement of peer reviewers. In both Sweden and Denmark, a gender-specific difference in the success rate of attainment of postdoctoral positions and research grants has been documented, to the detriment of women. Recently, the Massachusetts Institute of Technology admitted to having given fewer resources to female than male scientists of equal seniority (see <http://web.mit.edu/fnl/women/women>).

In *Why So Slow? The Advancement of Women*, Virginia Valian has compiled studies on the slow progress of women in professional life, its possible causes and potential remedies. Her theory is that a continuous accumulation of small advantages for men and small disadvantages for women operate insidiously, resulting in very different career opportunities for the sexes. A multitude of cited studies demonstrate that evaluators as a rule overestimate the performance of men and underestimate that of women.

Valian didactically explains the roots of prejudice and stereotyping. The simple belief that two groups of people are different

will taint the evaluation of an individual from either group, impeding a clear-headed evaluation of that person's merits. Stereotyping is a basic human inclination, which probably evolved, according to Valian, to facilitate speedy judgement — if one sees a tiger, it is better to have prejudiced ideas on the aggressiveness of that tiger, based on previous encounters with other tigers or hearsay, than to wait and see how this particular tiger behaves. Girls and boys, men and women, are perceived and treated differently even in cases where the two sexes behave in exactly the same way. Valian calls this stereotyping "gender schemata", and shows that women as well as men firmly believe in psychological differences between the sexes, when, in fact, no psychological tests have shown any such differences.

Valian's reasoning is straightforward and easy to follow, with one exception. Although she begins by rightly pointing out that women and men are alike in all measurable psychological variables, she contradicts herself by referring to studies where women are shown to be more 'feminine' than men, who instead more often are 'masculine'. What kind of traits does Valian allude to? The 'masculinity-femininity scale' was constructed by American psychologists based on people's opinions of what constituted typically feminine and masculine traits. Based on which traits men and women picked from these lists to describe themselves, they were defined as being masculine or feminine. Here, academic psychology has achieved a perfect circular reasoning in order to preserve outdated gender stereotypes. This is not Valian's fault — but the book would have done better without it.

Valian argues that women and men are equally biased against female performance. This is indeed true of experimental psychology studies performed in the 1970s and early 1980s, but some more recent studies indicate that women are adopting a more gender-neutral evaluation behaviour while men continue to underestimate female achievement. Most likely, in the past 20 years, women have become aware of gender inequalities in society and have tried to correct their behaviour. This point is of some importance, because if women are less biased against female achievement than are men, there is urgent reason to increase the number of women evaluators.

This book should be read by anyone interested in the field of gender equality, but, more importantly, by those who are not interested in such issues. The ideal target group should be people in power who live under the illusion that they 'always treat men and women alike'.

Christine Wennerås is in the Department of Medical Microbiology and Immunology, and Agnes Wold is in the Department of Clinical Immunology, Göteborg University, Guldhedsgatan 10, SE-413 46 Göteborg, Sweden.

Science in culture

Hay-on-Wye Festival of Literature and the Arts

(28 May–6 June 1999)

John Maddox

Reports to the contrary notwithstanding, popular enthusiasm for science is alive, well and brimming over in this improbable setting. Hay-on-Wye is an old Welsh border town (where the border is that with England, now pushed further east under new parliamentary arrangements). The town was revived and put on the map in the 1960s by one Richard Booth, whose entrepreneurship has given it a large network of second-hand bookshops. Quite independently, the town has been for the past decade the site of a bucolic literary festival, held in huge tents scattered around the elementary school, with views of rolling fields and hills rendered intense green by the heavy and persistent rain familiar in these parts.

Science has by accident infiltrated the standard programme of a literary festival. If it makes sense to celebrate the latest Doris Lessing or Edna O'Brien book by asking the authors to come and talk, why omit Stephen J. Gould, Stephen Hawking and Stephen Pinker? (This festival also embraces politics — witness the regular visits by the handful of literate British parliamentarians and by luminaries such as F. W. de Klerk, quondam president of South Africa.)

This year's science offerings outdid all previous years, which is partly a measure of the increased volume of popular-science books on the market and partly a deliberate decision by the organizers to give the audience more of what it seems to clamour for. This year, the geneticist Steve Jones (from University College London) almost filled the 1,000 seats in the largest of the tents with a talk on the question of whether human evolution has now come to a halt. But there were a dozen other science talks spanning cosmology (by Paul Davies from Adelaide, John Gribbin — an old *Nature* hand — and Chandra Wickramasinghe, the iconoclast from Cardiff) and the question of where and what is mind (by Bristol psychologist Richard Gregory, journalist Rita Carter, neurologist Christopher Frith and Ian M. Glynn, the now retired professor of physiology at Cambridge).

Studied informality is the order of the day for presenters. Steve Jones, a regular visitor, this year abandoned his customary sweater and slacks for a dark double-breasted suit and gleaming white shirt (but no tie), but

that was only illusion: in no time at all, the jacket was thrown over the back of a chair and Jones was back in his element. When the rain thundered on the roof of the tent, Jones was able to send the audience into gales of laughter simply by turning his eyes towards the heavens.

Questions at Hay are thoughtful and intelligent, if usually prefaced by "I am no scientist". Mercifully, the audience seems to have given up asking every speaker whether he or she believes in God (maybe the answers were too uniform). But this year Jones ducked the question of whether, if evolution "by means of natural selection" had now halted, people might themselves take a hand in the process by genetic manipulation. These occasions need a chairman.

Meanwhile, why are huge crowds flocking to science talks at Hay? Much of the audience seems to consist of schoolteachers occupying the week-long break that British schools enjoy at the end of May. The festival director, Peter Florence, says that those who come to Hay are hungry for learning of all kinds, not just for refresher courses in their own fields. Whatever the explanation, on this year's showing, it is an occasion at which authors whose books have been ignored by reviewers in the fashionable press can be sure of an appreciative audience and can even be flattered by the chance to sign books in a crowded tent with rain hammering on the roof.

John Maddox is at 9 Pitt Street, London W8 4NY, UK.

● Next year's festival runs from 27 May to 4 June.



Hay-on-Wye: home to bookshops owned by Richard Booth (left), Hay's self-proclaimed king.

CORBIS/DAVE G. HOUSER
PAUL ROGERS

A record of atmospheric halocarbons during the twentieth century from polar firn air

James H. Butler^{*}, Mark Battle^{†‡}, Michael L. Bender^{†‡}, Stephen A. Montzka^{*}, Andrew D. Clarke[§], Eric S. Saltzman^{||}, Cara M. Sucher^{†‡}, Jeffrey P. Severinghaus^{†‡} & James W. Elkins^{*}

^{*} National Oceanic and Atmospheric Administration, Climate Monitoring and Diagnostics Laboratory, Boulder, Colorado 80303, USA

[†] Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island 02882, USA

[§] Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado 80309, USA

^{||} Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida 33149, USA

Measurements of trace gases in air trapped in polar firn (unconsolidated snow) demonstrate that natural sources of chlorofluorocarbons, halons, persistent chlorocarbon solvents and sulphur hexafluoride to the atmosphere are minimal or non-existent. Atmospheric concentrations of these gases, reconstructed back to the late nineteenth century, are consistent with atmospheric histories derived from anthropogenic emission rates and known atmospheric lifetimes. The measurements confirm the predominance of human activity in the atmospheric budget of organic chlorine, and allow the estimation of atmospheric histories of halogenated gases of combined anthropogenic and natural origin. The pre-twentieth-century burden of methyl chloride was close to that at present, while the burden of methyl bromide was probably over half of today's value.

Chlorofluorocarbons (CFCs) and other halocarbons have been used as refrigerants, propellants and solvents since the early to middle twentieth century. Both leakage and direct emission, in combination with long atmospheric lifetimes, have resulted in a steady accumulation of these gases in the atmosphere, allowing them to increase the burden of atmospheric chlorine (Cl) from $\sim 0.5 \text{ nmol Cl mol}^{-1}$ (that is, 0.5 parts per billion, p.p.b.) to a peak of over $3.5 \text{ nmol Cl mol}^{-1}$ by 1992–94^{1,2}. Suggestions that a build-up of CFCs and long-lived halocarbons in the atmosphere could result in the depletion of stratospheric ozone prompted the need for systematic measurements³, but only in the mid-1970s were routine atmospheric monitoring programmes begun^{4,5}. Industrial production and emission data, together with atmospheric lifetimes, have been used to estimate the atmospheric burden of CFCs before their detection and measurement^{6,7}. Until now, however, no data have been available for testing these early estimates.

Natural sources have been suggested for many of the halocarbons involved in the depletion of stratospheric ozone⁸. Atmospheric methyl chloride (CH_3Cl), for example, is believed to be predominantly of natural origin, although it does have a few anthropogenic sources and there appear to be unidentified sources of this gas⁹. The division of methyl bromide (CH_3Br) emissions into natural and anthropogenic components has generated considerable scientific and regulatory controversy during recent years^{10,11}. Some studies have suggested that natural sources may be significant for atmospheric CCl_4 (ref. 12). Detection of CFCs in air sampled near volcanoes has led to the suggestion that volcanism might be a significant source of these gases in air¹³. The general absence of CFCs in the ocean's deep waters virtually precludes this possibility, but no reliable measurements in air taken before the onset of anthropogenic emissions have been made to date. Air samples from ice cores have been too small to obtain the precision necessary to address these questions for halocarbons. Previous collections of firn air

either were contaminated during sampling¹⁴ or extended back no more than a few years¹⁵.

Here we use samples of air collected from firn (unconsolidated snow; Table 1) at the South Pole and Siple Dome in Antarctica and at Tunu in Greenland to obtain high-precision measurements spanning the full history of anthropogenic emission of most halocarbons (data reported here are available; see Supplementary Information). Samples of firn air from all three sites were analysed for CCl_3F (CFC-11), CCl_2F_2 (CFC-12), CH_3CCl_3 , CCl_4 , N_2O , SF_6 , CBrF_3 (halon H-1301), CBrClF_2 (halon H-1211), CH_3Br and CH_3Cl . Newly introduced hydrofluorocarbons (for example, $\text{CF}_3\text{CH}_2\text{F}$ or HFC-134a), hydrochlorofluorocarbons (such as $\text{CH}_3\text{CCl}_2\text{F}$ or HCFC-141b, CH_3CClF_2 or HCFC-142), an additional CFC ($\text{CCl}_2\text{FCClF}_2$ or CFC-113), and other partially halogenated methanes were measured in Tunu and Siple Dome samples only. (CHClF_2 , or HCFC-22, showed significant signs of contamination in the upper samples, so have been omitted from this analysis.) Results for N_2O , CO_2 , CH_4 , O_2/N_2 and ^{15}N at the South Pole already have been published¹⁶. Depth profiles from these sites have been presented in two NOAA/CMDL data reports^{17,18}. Owing to physical

Table 1 General properties of firn at sampling sites

Site	South Pole, Antarctica	Siple Dome, Antarctica	Tunu, Greenland
Location	90° S	81° 40' S, 148° 49' W	78° 01' N, 33° 59' E
Elevation (m)	2,841	~600	~2,400
Sampling date	Jan. 1995	Dec. 1996	Apr. 1996
Lock-in zone depths (m)	114–122	48–56	58–68
Annual snow accumulation (cm yr^{-1} ice equivalent)	8	15	10
Annual mean temperature (°C)	–49.4	–25.4†	–29†
'CO ₂ age' at bottom*	1903	1953	1929†

* CO₂ age is defined as the year the atmospheric CO₂ value³³ matched the CO₂ concentration in the deepest sample. This is a mean age for CO₂ in the bottom air, which has undergone considerable mixing. Mean ages of halocarbons are greater because of their lower diffusivities. Once in the lock-in zone, they then age at the same rate. This value is given to illustrate differences among the sampling sites; firn air at the South Pole records atmospheric history over longer time than the other two sites.

† This value is approximate owing to the unknown evolution of the interhemispheric CO₂ before 1972.

‡ Temperatures from measurements made at the site at the time of sampling.

‡ Present addresses: Department of Geosciences, Princeton University, Princeton, New Jersey 08544, USA (M.B., M.L.B.); ‡ National Oceanic and Atmospheric Administration, Climate Monitoring and Diagnostics Laboratory, Boulder, Colorado 80303, USA (A.D.C.); US Global Change Research Program, Washington DC 20024, USA (C.M.S.); Scripps Institution of Oceanography, La Jolla, California 92037, USA (J.P.S.).

effects, and in some instances chemical effects, the composition of firn air at any given depth does not correspond exactly to the atmospheric composition at some time in the past. Nevertheless, because the behaviour of each gas in the firn is governed primarily by its own diffusivity and molecular mass, we are able to estimate histories of both anthropogenic and naturally occurring halocarbons during a time of rapid population, agricultural and industrial growth.

Concentration–depth profiles of trace gases

Firn is a porous matrix in which constituent gases diffuse at rates governed by both the structure of the firn and the nature of the species in question. Profiles of $\delta^{15}\text{N}$ values of N_2 versus depth in the firn (Fig. 1) reflect the processes governing transport and storage of air^{19,20}. Below the influence of seasonal temperature gradients and other surface-related effects (roughly 30 m), $\delta^{15}\text{N}$ of N_2 increases linearly with depth, owing to gravitational settling. The linearity indicates a porous column in which transport is governed predominantly by molecular diffusion^{21,22}. The interval of constant $\delta^{15}\text{N}$ of N_2 at the base of the firn shows the lock-in zone²³, where dense winter layers substantially impede vertical diffusion, while summer layers retain enough open porosity to allow sample collection. Gases within the porous column atop the lock-in zone communicate with the atmosphere through diffusion, creating a smoothed record of atmospheric changes. In contrast, vertical diffusion is considerably restricted within the lock-in zone, which tends to isolate this air from the diffusive column. Thus, air in the lock-in zone ages at approximately the same rate as the surrounding ice. At Siple Dome and Tunu, where the diffusive column is relatively short and smoothing is minimal, the composition of the overlying atmosphere is recorded in the lock-in zone without extensive diffusive mixing. At the South Pole, the deeper diffusive column significantly smooths the atmospheric history, but also allows for collection of very old samples¹⁶. By comparing results from these sites with different diffusive structures, we are able to reconstruct robust atmospheric histories.

Variations in concentration with depth reflect both long-term trends and recent perturbations of trace gases in the atmosphere (Figs 2–4). Concentrations of compounds known to have sig-

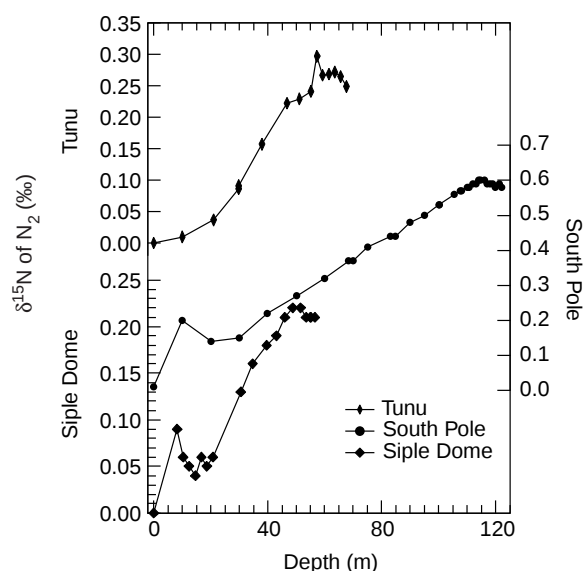


Figure 1 Measured depth profiles of $\delta^{15}\text{N}$ of N_2 . The results from the two boreholes at the South Pole are indistinguishable from one another. Enrichments of ^{15}N at shallow depths reflect thermal fractionation associated with seasonal temperature changes at the surface. $\delta^{15}\text{N}$ (in ‰) = $\{[(^{15}\text{N}/^{14}\text{N})_{\text{sample}} / (^{15}\text{N}/^{14}\text{N})_{\text{standard}}] - 1\} \times 10^3$; the standard used here for all sites is contemporary tropospheric air.

nificant anthropogenic sources over the past few decades (for example, CFCs, solvents and halons) decrease considerably with depth, where the firn air is older. Concentrations of CFC-11, CFC-12, CFC-113, H-1301, H-1211 and SF_6 all decreased to levels that did not differ significantly from zero by the bottom of the depth profiles (Fig. 2). This indicates that these gases were not present in the atmosphere during the early twentieth century and confirms their exclusively anthropogenic origin. Previously, the lowest values reported for CFC-11 and -12, also from firn air measurements, were 16 and 18 pmol mol^{-1} (that is, parts per trillion or p.p.t.)¹⁴. The earliest direct measurements of atmospheric CFC-11 and -12 date from the early to mid-1970s, with values of roughly 50 and 200 pmol mol^{-1} (refs 12, 24–26). CFC-113 has been measured in archived air from Cape Grim, Australia, as far back as 1978, for which a value of 13 pmol mol^{-1} was reported²⁷. Systematic SF_6 measurements began in 1978 at levels of 0.6 pmol mol^{-1} , or about 15% of the current atmospheric burden²⁸, although the lowest reported value was for 1976, at 0.24 pmol mol^{-1} (ref. 29). Early measurements of the halons are also available since 1976, with amounts of the order of 0.5 pmol mol^{-1} (ref. 30).

The concentrations of two abundant chlorocarbons, CH_3CCl_3 and CCl_4 , also decreased rapidly with depth (Fig. 3). Both of these gases reached concentrations near the bottom of the profiles that were at or near detection limits, suggesting that natural sources are insignificant and that both gases were essentially absent in the pre-twentieth-century atmosphere. The detection limit for CH_3CCl_3 was $\sim 1.5 \text{ pmol mol}^{-1}$, although for CCl_4 it was $\sim 5 \text{ pmol mol}^{-1}$ because of a co-eluting peak on the gas chromatography/electron-capture detection (GC/ECD) system (see Supplementary Information). Thus natural sources, if they are significant at all, could not be responsible for more than 5 pmol mol^{-1} of CCl_4 in the atmosphere. Of course, this requires that CCl_4 is stable in firn air, a question that is raised in our attempts to reconcile an atmospheric history of CCl_4

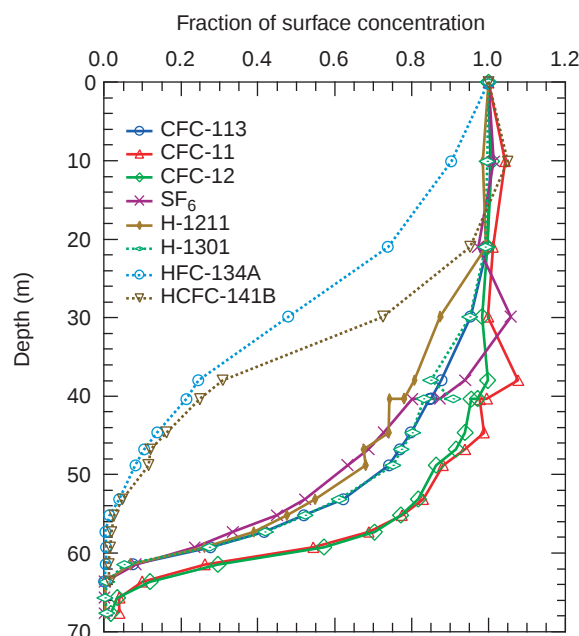


Figure 2 Measured depth profiles of anthropogenic gases with 'known' emissions (Tunu, Greenland). These profiles are similar to those obtained at South Pole and Siple Dome. Depth profiles at all sites show the CFCs, halons and SF_6 penetrating deeply into the firn and dropping to zero or near-zero at the bottom, which implies no significant sources before the introduction and use of these gases by humans in the middle to late twentieth century. In general, gases emitted for longer periods and with higher diffusivities penetrate most deeply. The atmospheric increases of HCFC-141b and HFC-134a (defined in the text) are recorded only in the diffusive zone of the firn, as these gases have not been present long enough in the atmosphere to become locked in at depth.

from both Siple Dome and South Pole profiles. Although absolute calibration was a concern for some of the early real-time measurements of these two gases^{12,24}, values from the late 1970s were reported at around 60 pmol mol⁻¹ for CH₃CCl₃ and 90 pmol mol⁻¹ for CCl₄, after normalizing to recent calibration scales^{2,31}. Also apparent in the data from Siple Dome (Fig. 3) and Tunu is the 1991–93 turnaround of CH₃CCl₃ in the atmosphere that resulted from decreased anthropogenic emissions^{1,31}. Those HCFCs and HFCs that have been used only recently appear exclusively in the upper layers of the firn, as they have not been present in the atmosphere long enough to have been isolated in the lock-in zone¹ at concentrations above our detection limits (Fig. 2).

Concentrations of CH₃Cl are significant in the deepest samples, and increase gradually toward the surface of the firn (Fig. 3). This is evidence that this gas has both natural and anthropogenic sources or at least sources that existed before the twentieth century. The concentrations of CH₃Cl in samples collected between 20 m and 45 m depth are similar to the mean annual concentrations at Cape Grim, Tasmania, in recent years¹⁸. The decrease of 15–20 pmol mol⁻¹ observed in the upper 10–12 m of the firn is consistent with the seasonal variability observed for CH₃Cl in the Southern Hemisphere, as lower mixing ratios are observed in summer when atmospheric loss processes outweigh sources. Concentrations between 20 and 45 m depth are about 10% higher than at the bottom of the profile. If we assume that CH₃Cl is not being degraded slowly over time in the firn, these results indicate that the

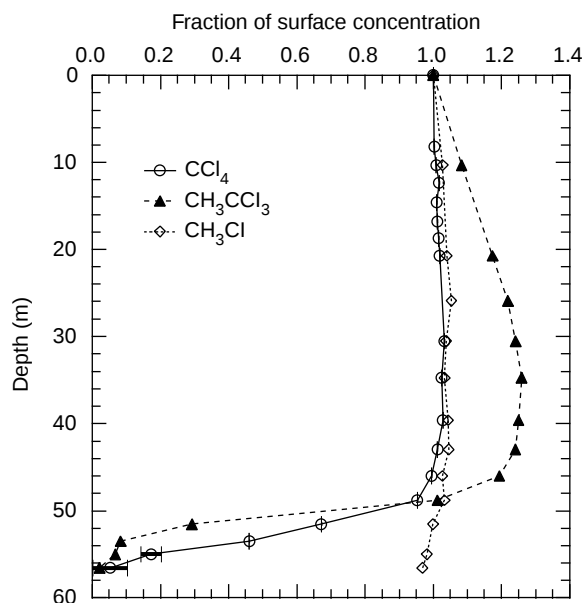


Figure 3 Measured depth profiles of chlorocarbons (Siple Dome, Antarctica). Profiles at Tunu and the South Pole were similar. Atmospheric histories of these gases are poorly known. Natural sources have been suggested for each of these species, but these profiles show that natural or pre-twentieth-century sources are large only for CH₃Cl. CCl₄ concentrations are significant at the bottom of this profile, which nominally represents the early 1950s. At Tunu, where the mean age of the bottom samples is in the early 1930s, and at South Pole, where the mean age dates from the early 1900s, the CCl₄ concentration near the bottom was essentially zero. The detection limit for CCl₄, however, was ~5 pmol mol⁻¹, owing to an interfering peak in the GC/ECD system. Traces of CH₃CCl₃ were found in some of the deepest samples at all sites, but these may be artefacts associated with filling high-pressure steel flasks at the bottom of the hole or storage in the glass flasks. The detection limit for CH₃CCl₃ was ~1.5 pmol mol⁻¹. Note also that the 1993 peak in atmospheric CH₃CCl₃ in the Southern Hemisphere^{1,31} appears in the firn near 40 m depth at this site; this peak (1991–92 in the Northern Hemisphere) also appeared in the Tunu depth profiles (see Supplementary Information).

atmospheric mole fraction of CH₃Cl has increased by about 40 pmol mol⁻¹ over the past century.

The firm data for CH₃Br are more difficult to interpret. At the two Antarctic sites, CH₃Br concentrations at the top of the firn air profiles are 20–25% higher than at the bottom, implying a significant pre-twentieth-century source and an increase in the atmospheric burden of this gas over the past 50–100 years (Fig. 4). Unlike CH₃Cl, however, the observed difference of 5–10% between the surface samples and those below 10 m is not explained by seasonally varying mixing ratios; no significant seasonality is observed in the atmosphere at sites in the Southern Hemisphere³². This anomaly remains a mystery, as it was present at both Antarctic sampling sites and appears not to be a sampling or analytical artefact. More disturbing, however, are the samples from Tunu. At this seasonally warm and coastally influenced Greenland site, the concentration of CH₃Br was very high near the bottom of the profile, reaching mixing ratios of nearly 50 pmol mol⁻¹ at the firn–ice transition (Fig. 4). We are confident that this increase in concentration at depth is not an artefact of sample collection, storage or analysis. CH₃Br–depth curves were essentially identical in our two Tunu firn-air profiles, although the second profile did not extend as deeply as the first. Tests of the sampling equipment showed no evidence whatsoever of contamination by CH₃Br (CFCs were present in our equipment, but our sampling design and procedure still allowed for concentrations of zero or nearly zero in deep samples.) Samples in steel and glass flasks yielded similar profiles for CH₃Br and CH₃Cl, and chromatographic results, obtained by both electron-capture or mass spectrometry, were independent of analytical technique. This leads us to believe that the observed high values (up to 48 pmol mol⁻¹) for CH₃Br in the firn at Tunu are real, although not necessarily of atmospheric origin. Concentrations of other marine, biogenic gases (such as CH₂Br₂ and CH₃I) also increased at depth in the Tunu firn, although the anomalies were smaller than for CH₃Br.

Interpreting concentration–depth data

Obtaining a representative atmospheric history of any gas from a diffusion model and measurements of firn air requires some knowledge of the diffusivity of gases in the firn. This can be derived from

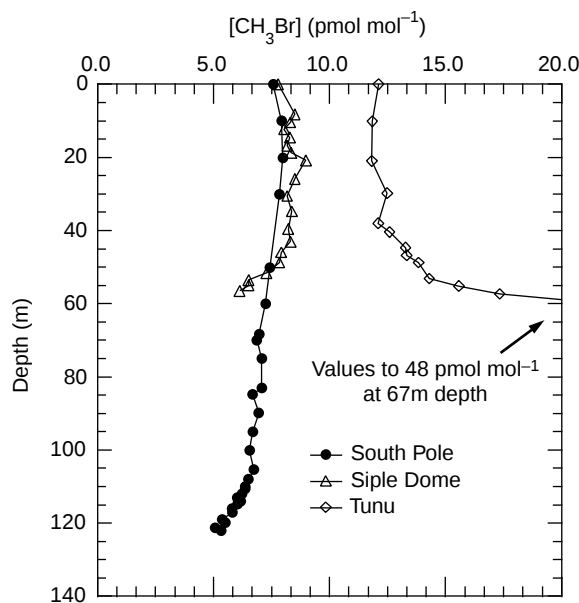


Figure 4 Measured depth profiles of CH₃Br at all three sites. Whereas the profiles from South Pole and Siple Dome imply a similar atmospheric history involving both natural and anthropogenic sources, the data from Tunu are in direct contrast to the Antarctic observations and indicate that atmospheric CH₃Br signals are not necessarily conserved in firn air.

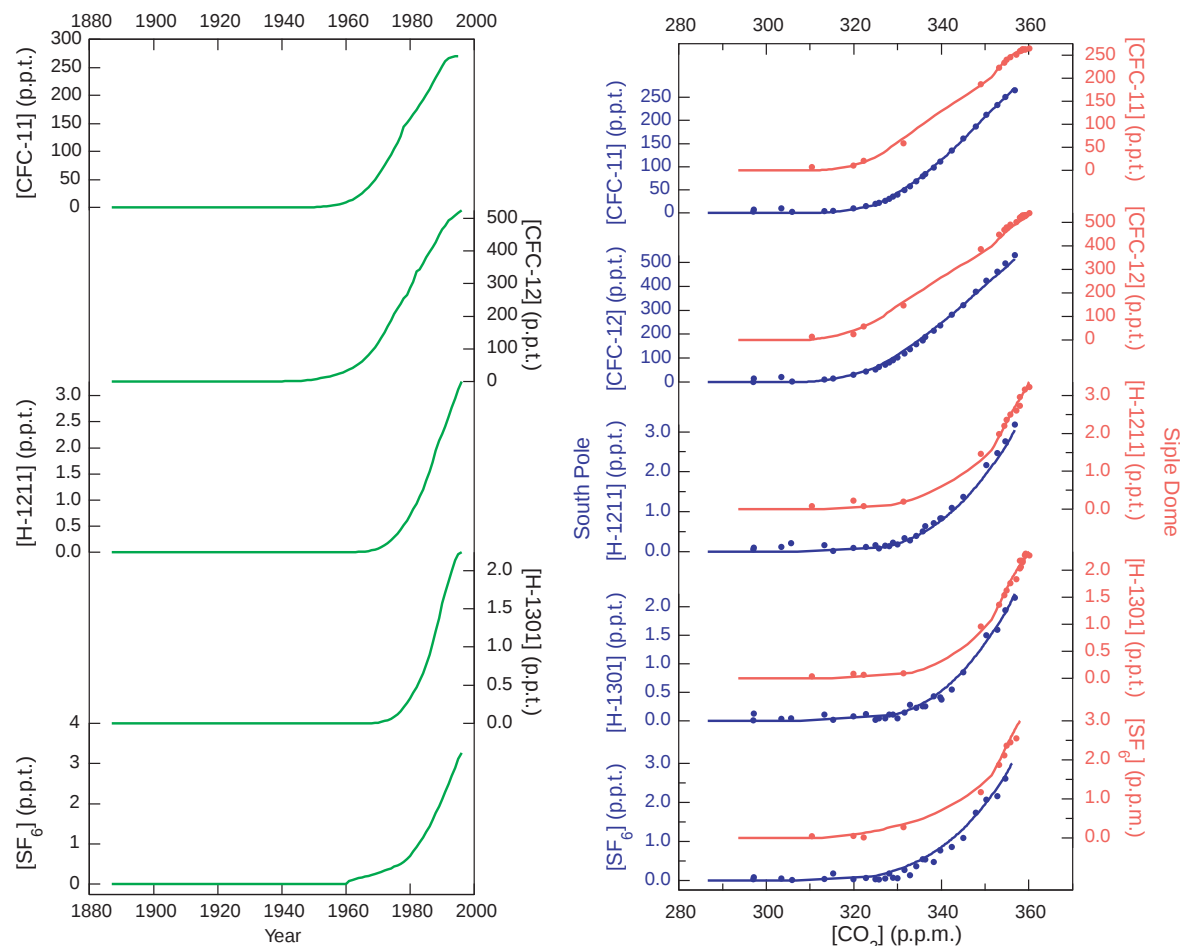


Figure 5 Histories of CFCs, halons and SF_6 in the Southern Hemisphere. Data are for Antarctic firn air only. Modelled values are shown as lines and measured mole fractions are shown as points (p.p.t., pmol mol^{-1}). Assumed Southern Hemispheric histories (green lines) for each of these gases are obtained from combinations of industrial emission models and real-time measurements^{1,35–37} (left panel). These histories are incorporated into the firn at both sites with the one-dimensional firn diffusion model. Model results are then plotted as a function of CO_2 in firn air (right panel) for the South Pole (blue) and Siple Dome (red). (Note offset scales for each site.) Diffusivities used in these models are calculated from Wilke and Lee⁴⁸, except for CFC-11 which was lowered by 10%, but still remains

within uncertainties. Lifetimes used for the gases are those given in the 1994 Scientific Assessment of Ozone Depletion³⁸, except for H-1211, which was lowered to 11 years to assure concordance with current atmospheric measurements³⁷. The good agreement between model results and measurements at these two sites, which have different firn depths and diffusion profiles, indicates that the assumed histories are consistent with concentrations observed in the firn. The detection limit for SF_6 in the South Pole samples was $0.3 \text{ pmol mol}^{-1}$, a value that was much improved for analysis of samples from the other two sites.

density–porosity–diffusivity relationships¹⁹. These relationships, however, are poorly constrained, so we prefer to calibrate diffusivity with the measured atmospheric history of CO_2 deduced from real-time flask samples and the Law Dome ice core³³. We refrain from modelling the Tunu profile because of the extensive convective zone and the possibility that the firn-air CO_2 profile at that site is erroneously offset by covariance between seasonal convective mixing and seasonal variations in atmospheric CO_2 ('rectifier' effects). To account for the effects of diffusion in the firn on each species at the South Pole and Siple Dome, we use the simple one-dimensional firn diffusion model described in detail by Battle *et al.*¹⁶. As in that investigation, we infer a site-specific diffusivity–depth relationship by adjusting diffusivity until the model reproduces the observed CO_2 firn-air profile when driven by the atmospheric CO_2 history. We then calculate halocarbon–depth profiles using this adjusted diffusivity–depth profile in the diffusion model. However, we examine these as halocarbon– CO_2 rather than halocarbon–depth profiles; the former are much less sensitive to errors in the inferred profiles of diffusivity versus depth. Our results remain sensitive to errors in the diffusivities of halocarbons relative to the diffusivity of CO_2 .

We apply the firn diffusion model with one of two goals, depending on the halocarbon: (1) to test an independently derived atmospheric history, or (2) to infer an atmospheric record for species with uncertain or unknown histories. The first approach is suitable for gases such as the CFCs, halons and SF_6 , for which atmospheric histories have been derived from emission estimates and real-time measurements. We use these histories and the firn-air diffusion model to predict concentration– CO_2 profiles for these gases and to compare the predictions with observations in the Antarctic firn. The second approach is used for those chlorocarbons and methyl halides for which there is insufficient information to estimate or calculate complete atmospheric histories. In this approach we choose one or more suitably general mathematical descriptions of concentration as a function of time, with consideration given to recent direct atmospheric measurements as well as firn-air data. The free parameters of the mathematical expressions are then adjusted using inverse methods to optimize agreement between measured and modelled concentration– CO_2 profiles at one or both sites.

We cannot eliminate the possibility that these gases are very slowly produced or degraded in the firn. However, we have strong

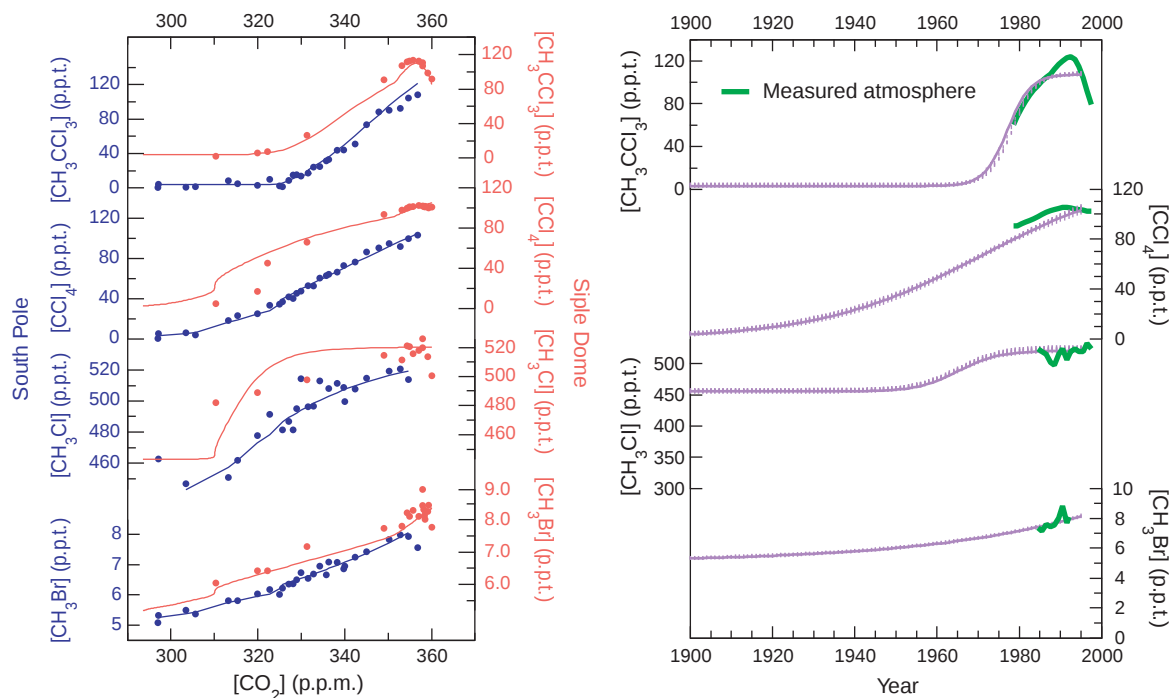


Figure 6 Histories in the Southern Hemisphere of halocarbons with uncertain emission records. Modelled (lines) and measured (points) mole fractions of CCl_4 , CH_3CCl_3 , CH_3Cl and CH_3Br are shown plotted against CO_2 in firn air (left panel) for the South Pole (blue) and Siple Dome (red) (note offset scales for each site). The model results represent gas mole fractions (p.p.t., pmol mol^{-1}) predicted in the firn based on the atmospheric histories shown in the right panel (purple). These histories are inferred by optimizing the data-model agreement for the South Pole (blue lines) and then checking this for agreement with data from Siple Dome (red lines). Also shown are real-time atmospheric measurements (in green) from the

ALE/GAGE/AGAGE network^{2,50,51} and from Khalil *et al.*^{9,52}, normalized to NOAA/CMDL South Pole measurements¹. In the inversion model, real-time atmospheric measurements have been used only to adjust the diffusivities⁴⁹ of CH_3Br (–20%) and CH_3CCl_3 (–20%). The shapes of the inversion results are independent of these measurements. Atmospheric concentration histories shown here follow sigmoid curves, with parameters determined by the firn data. The use of other functional forms for the atmospheric histories yields indistinguishable results. Errors (shown as purple bars) correspond to a 68% confidence envelope.

evidence that such production or degradation is not significant for most of them. Agreement of emission records with the firn model for those gases with available data provides one level of assurance. Consistent histories inferred from measurements at Siple Dome and the South Pole provide additional assurance; Siple Dome and the South Pole are sites sufficiently different in firm properties that any physical, biological or chemical effects on the gas concentrations should be expressed differently at these two sites.

Tests of independently derived atmospheric histories

Concentration– CO_2 profiles predicted for CFCs, halons and SF_6 based on emission histories^{6,34}, real-time measurements^{1,35–37} and firn-air diffusion modelling are in good agreement with firn-air data (Fig. 5). This agreement, together with zero or near-zero concentrations in the oldest samples, is consistent with solely anthropogenic sources of these gases. It also suggests that the emission models used^{7,36} yield reasonably accurate estimates of the early atmospheric histories.

The one significantly inconsistent species is halon H-1211. Both modern measurements and firn results suggest either a shorter lifetime than that given solely by atmospheric chemistry models (11 compared to 20 years) or lower emissions^{37,38}. The atmospheric curve and model output in Fig. 5 incorporate the 11-year lifetime. For the CFCs, for which significant emissions began in the 1940s, analyses of firn air provide data for long periods where few or no real-time measurements are available³⁶. For the halons, with significant emissions beginning in the 1960s, real-time measurements span a larger fraction of their atmospheric history, although calibrations of earlier measurements are less certain^{30,37}. For SF_6 , with significant emissions beginning in the 1970s, the concentration history is derived almost entirely from direct atmospheric

measurements^{28,35}, and thus is insensitive to errors in estimates of emissions or lifetimes.

Inferred atmospheric records of chlorocarbons

CCl_4 appears to be derived solely from anthropogenic sources, but this conclusion must be tempered by the analytical uncertainty of $\pm 5 \text{ pmol mol}^{-1}$ for the lowest-concentration samples, owing to an interfering peak in the GC/ECD system. This inference is based primarily upon the longer South Pole firn-air record (Fig. 6), although the concentration of CCl_4 also does not differ from zero at the bottom of the Tunu profile. The firn data do show that CCl_4 has been present in the atmosphere for a longer time than the CFCs, owing to its earlier industrial use. This earlier use also appears in the deep ocean record where CCl_4 is present in the relative absence of CFCs³⁹. Although our data from Siple Dome suggest that, before 1960, CCl_4 was substantially lower in concentration than predicted from estimated production and emissions⁴⁰, the data from the South Pole agree quite well with the emission prediction. Emission estimates for CCl_4 , however, are much less certain than for the CFCs or halons. The lowest value at Siple Dome, corresponding to the late 1940s or early 1950s, was 4–7 pmol mol^{-1} , but the South Pole data suggest a value of about 35 pmol mol^{-1} for 1950. In an early analysis of CCl_4 emissions, Galbally⁴¹ used a lifetime of 38 years to derive an expected atmospheric mixing ratio of 40 pmol mol^{-1} . More recently, and using a lifetime of 35 years, Walker *et al.*⁴⁰ suggest a value of about 33 pmol mol^{-1} for CCl_4 in the Southern Hemisphere in 1950. Though uncertainties are large, both of these estimates from emission records are reasonably consistent with our South Pole data since the 1950s, yet are lower for earlier years. Unfortunately, the inconsistency between Siple Dome and South Pole data suggests that CCl_4 is, to some extent, either consumed or absorbed in the

firn; thus, we cannot make an unambiguous reconstruction of its concentration history.

Although the derived atmospheric record of CH_3CCl_3 is consistent at the two Antarctic sites, which suggests that this gas is stable in firn air, two data points in the lower part of the South Pole record do show measurable amounts of CH_3CCl_3 . These, however, are most probably caused by contamination. The lowest value at each of the three sites, representing years from the 1890s to the 1940s, was $<2 \text{ pmol mol}^{-1}$ ($<2\%$ of maximum atmospheric levels). Such concentrations are very near or below our limit of detection ($3\text{--}5 \sigma$; see Supplementary Information) and probably result from low-level contamination in the sampling apparatus. Note also that, although the turnaround in atmospheric CH_3CCl_3 was recorded in the firn at Tunu (Fig. 3), this was not the case at South Pole (Fig. 6). This is because the South Pole was sampled only shortly after the turnaround. There is a time lag of about one year in transporting signals from the Northern Hemisphere to the Southern Hemisphere and wind pumping causes some mixing in the uppermost part of the firn, so the signal would be suppressed or absent in the South Pole firn in early 1995.

Because the inferred temporal change in CH_3Cl is small relative to its variability in the firn air, a detailed history cannot be estimated reliably from these data (Fig. 6). Nevertheless, the difference of $5\text{--}10\%$ between bottom and mid-hole samples remains significant and suggests that the atmospheric burden of CH_3Cl has increased by about this amount during the past century. CH_3Cl is the most abundant chlorine-containing gas in the atmosphere and is emitted from the ocean⁴², from biomass burning⁴³, and by fungi⁴⁴. Its current atmospheric budget remains uncertain. Real-time measurements of atmospheric CH_3Cl extend back only to 1980⁹ and show wide seasonal oscillations. The firn data represent a much longer time span, and are compatible with the real-time measurements.

Methyl bromide

Methyl bromide measurements from the South Pole and Siple Dome suggest a simple atmospheric history, with a rate of growth increasing from $0.01 \text{ pmol mol}^{-1} \text{ yr}^{-1}$ in the early 1900s to $0.05\text{--}0.06$ (± 0.01) $\text{pmol mol}^{-1} \text{ yr}^{-1}$ (90% confidence limits) during the 1970s and 1980s. These two sites differ in snow accumulation rate, firn depth and temperature (Table 1). Thus the similarity of the two $\text{CH}_3\text{Br}\text{--CO}_2$ profiles suggests that the record in the firn is atmospheric in origin. Taken at face value, such a record implies that the concentration of CH_3Br was increasing slowly in the atmosphere through the first half of the century, perhaps as a function of increased biomass burning or some other global change. The upturn in the mid-1960s coincides with the onset of CH_3Br use as an agricultural fumigant, which suggests that the increase since the 1960s could indicate the expected reduction, about $15\text{--}25\%$, from the international ban on the use of this gas as a fumigant⁴⁵. However, the response of CH_3Br in the atmosphere to a ban on agricultural use is also dependent upon how other sources have changed since 1960. The budget of CH_3Br for the modern atmosphere, calculated from what is known of sources and sinks, is grossly out of balance. Based upon current understanding^{46,47}, there is a large source of this gas that has not been identified. From the atmospheric concentrations implied by the deepest samples at the South Pole, it appears that this 'missing' source was also present at the turn of the century. These results suggest that the unidentified source is separate from modern fumigation practices, or that the overall sink, calculated for CH_3Br from atmospheric reactions and loss to the oceans and soils, is currently overestimated.

The main problem in interpreting the CH_3Br data comes from the Tunu site, where concentrations of CH_3Br increased with depth (Fig. 4). The Tunu record does not appear to reflect past atmospheric changes, and it casts some doubt upon the utility of the Antarctic data. The simultaneous occurrence of CH_3Br at 50 pmol mol^{-1} in the Northern Hemisphere and 5 pmol mol^{-1} in

the Southern Hemisphere is highly improbable, as it would require an atmospheric lifetime of 0.1 year or less for CH_3Br . All evidence today points to a lifetime of $0.5\text{--}1.2$ years (ref. 47). Locally increased atmospheric concentrations are possible, but unlikely at these levels in remote areas. The most logical explanation is that CH_3Br has been injected into the firn air at Tunu, perhaps by desorption or *in situ* production. Tests with our firn diffusion model indicate that one can account for the anomalous enrichments only by invoking net injection of CH_3Br at or near the firn–ice transition (see Supplementary Information). We know of no physical, chemical or biological mechanism that could easily explain this injection. Simple adsorption followed by subsequent release during densification would have injected CH_3Br into the firn air at all three sites. Abiotic chemical transformation at such cold temperatures (-45°C) cannot be ruled out, although attempts to identify possible source materials are speculative. Tunu has significant marine influence and is located in an area affected by arctic haze, which is known to contain significant amounts of bromine⁴⁸. There is no evidence to date suggesting that organisms grow at the low temperature of the firn, but it is possible that loosely bound material could be slowly released from enzymes or cell surfaces. Whatever the source of the anomalous CH_3Br , we cannot prove it to be absent in Antarctica. However, if the process is driven at the firn–ice transition zone, as it most probably was at Tunu, then it could not have been significant at the South Pole or at Siple Dome, which yield observed profiles of concentration decreasing monotonically with depth to the bottom of the holes. Even a very low net production at the lock-in zone would have changed the shape of the profiles noticeably, and also would have elevated the bottom concentrations substantially. Competing, unknown production and degradation processes throughout the firn might have altered the firn-air concentrations of CH_3Br , but the probability is low that these processes would produce similar profiles at the two physically dissimilar Antarctic sites.

Summary

Antarctic and Greenland firn can provide reliable archives of many atmospheric halocarbons dating back to the late 1800s, extending records of gases that have been measured in real time only during the past two decades. Robust records of CFCs, halons and SF_6 extracted from the measurements of firn air agree with real-time estimates in the later parts of the profiles, corroborate emission models for some species and are useful in evaluating lifetimes and emission estimates for others. Although there are still important uncertainties, our data refine halocarbon histories and give experimental evidence about their pre-anthropogenic concentrations. It appears that for many species the cold, dry air in the firn preserves the gases. Except for CH_3Cl and CH_3Br , which have significant natural sources, all of the halocarbons studied appear to have been derived entirely from emissions during the twentieth century. CH_3Br and perhaps CCl_4 are, for unknown reasons, not always conserved in firn air, and other reactive halocarbons could conceivably behave similarly. □

Methods

Two holes were drilled at each site. At the South Pole, both boreholes reached the firn–ice transition; at Tunu and Siple Dome the second holes ended in the diffusive zone. We only consider data from the first hole at Siple Dome for our model because of inconsistencies in the sampling protocol. Low-pressure samples ($\sim 120 \text{ kPa}$) were collected into duplicate, 2.5-litre glass flasks at all three sites. Higher-pressure samples ($\sim 380 \text{ kPa}$) were also collected into single or duplicate 2.4-litre stainless-steel flasks at Tunu and Siple Dome. Details on drilling and flask filling can be found in Battle *et al.*¹⁶. Glass flasks were filled at all three sites with a pump (MB-158, Metal Bellows, Senior Flexonics, Sharon, Massachusetts) attached to nylon or Dekabon tubing, and analysed by GC/ECD for 11 trace gases. CO_2 and $\delta^{15}\text{N}$ of N_2 also were measured in air from the glass flasks. Surface air was excluded from the borehole by inflating a

Received 18 December 1998; accepted 28 April 1999.

- Supplementary information** is available on *Nature's* World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of *Nature*.

Acknowledgements. We thank those responsible for drilling at South Pole (D. Giles and J. Kyne), Tunu (J. Kyne and B. Bergeron) and Siple Dome (E. Ramsey, J. Brown, and S. Root). We also thank R. Myers for preparation of standards used in this work. This research was supported by the NSF Office of Polar Programs, the National Institute of Global Environmental Change (NIGEC), the Atmospheric Chemistry Project of NOAA's Climate and Global Change Program, and the Methyl Bromide Global Coalition.

Correspondence and requests for materials should be addressed to J.H.B. (e-mail: jbutler@cmdl.noaa.gov). Data for gases at all sampling sites are available at <http://cmdl.noaa.gov/noah> or by anonymous ftp at <ftp://ftp.cmdl.noaa.gov/noah>.

Understanding the valency of rare earths from first-principles theory

P. Strange*, A. Svane†, W. M. Temmerman‡, Z. Szotek‡ & H. Winter§

* Theoretical Physics Group, School of Chemistry and Physics, Keele University, Keele, Staffordshire ST5 5BG, UK

† Institute of Physics and Astronomy, University of Aarhus, DK-8000 Aarhus C, Denmark

‡ Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK

§ INF, Forschungszentrum Karlsruhe GmbH, Postfach 3640, D-76021 Karlsruhe, Germany

The rare-earth metals have high magnetic moments and a diverse range of magnetic structures¹. Their magnetic properties are determined by the occupancy of the strongly localized 4f electronic shells, while the outer s-d electrons determine the bonding and other electronic properties². Most of the rare-earth atoms are divalent, but generally become trivalent in the metallic state. In some materials, the energy difference between these valence states is small and, by changing some external parameter (such as pressure), a transition from one to the other occurs. But the mechanism underlying this transition and the reason for the differing valence states are not well understood. Here we report

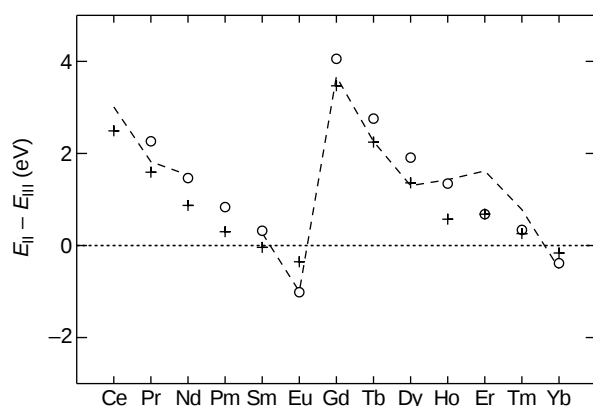


Figure 1 The energy difference (in eV) between the divalent and trivalent state of rare-earth materials. The dashed line shows the 'experimental' values for the rare-earth metals³. The open circles and the crosses show the calculated values for the rare-earth metals and the rare-earth sulphides, respectively. We see that the divalent and trivalent energy difference is large and positive at the beginning of the series, indicating that the trivalent state is well favoured. The energy differences then fall sharply and become negative for metallic Eu and for SmS and EuS, which means that the divalent state is the more stable one. There is a large discontinuous jump at Gd and GdS which are trivalent and in the latter half of the rare-earth series the values fall fairly steeply again, becoming negative at Yb for both the metal and the sulphide. As the divalent-trivalent energy differences calculated *ab initio* were too negative, they were uniformly shifted by 43 mRy to agree with the observed valence transition pressure of 6 kbar in SmS (ref. 7). This shift also fixes the energy differences for all the other rare-earth materials. In particular, for EuS we find a transition to the trivalent state to occur at a pressure of 184 kbar. Experimentally, the optical reflection spectrum⁸ shows an anomaly associated with a valence transition at ~160 kbar. For YbS we calculate a transition at 75 kbar, while similar experiments show movement towards the trivalent state at around 100 kbar (ref. 9). In the latter experiments the transition is shown clearly not to be to a purely trivalent state, but to be to some intermediate value of the valence, in good agreement with the present view.

first-principles electronic-structure calculations that enable us to determine both the valency and the lattice size as a function of atomic number, and hence understand the valence transitions. We find that there are two types of *f* electrons: localized core-like *f* electrons that determine the valency, and delocalized band-like *f* electrons that are formed through hybridization with the s-d bands and which participate in bonding. The latter are found only in the trivalent systems; if their number exceeds a certain threshold, it becomes energetically favourable for these electrons to localize, causing a transition to a divalent ground state.

Here we report a systematic theoretical investigation of the rare-earth elements and their sulphides using *ab initio* electronic-structure methods. These go well beyond standard calculations in which the 4f shell is described by an atomic model, while an itinerant picture for the s-d electrons is implemented^{3,4}. We include a self-interaction correction (SIC) which removes the spurious interaction of each electron with itself that occurs in conventional band-structure theory⁵. Our approach has the advantage of describing both the bonding s-d electrons and the *f* electrons on an equal footing. The SIC has a negligible effect on the bonding s-d electrons, but is substantial for the *f* electrons⁶. Application of the SIC to the *f* electrons provides a definition of valency of the metallic rare-earth materials. Here we associate the valency with the number of states available for the electron to propagate through the solid, namely

$$N_{\text{valency}} = Z - N_{\text{core}} - N_{\text{SIC}}$$

where *Z* is the atomic number of the rare earth and *N*_{core} is the number of atomic core electrons. The quantity *N*_{SIC} is the number of

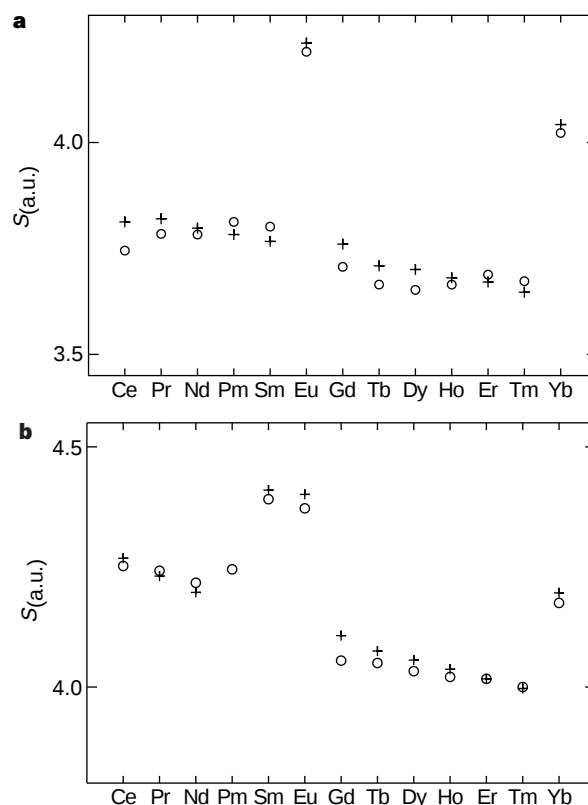


Figure 2 Wigner-Seitz radius as a function of atomic number for rare-earth materials. **a**, For the metals; **b**, for the sulphides. Crosses are the experimental values³ and circles are the calculated values. The Wigner-Seitz radius, *S*, is defined as the radius of a sphere with a volume equal to the actual crystal volume per formula unit. This eliminates effects due to crystal structure changes as we proceed across the periodic table; such effects are irrelevant to the present discussion. The abrupt jumps in Wigner-Seitz radius correspond to transitions between the divalent and trivalent states. (a.u., atomic units.)

f states for which the self-interaction has been removed and is determined so that N_{valency} equals 3 for trivalent, and 2 for divalent, systems. By performing *ab initio* SIC total energy calculations for all the rare-earth metals and their sulphides in the divalent and trivalent states we can determine their energy difference, together with the corresponding equilibrium volumes, and relate them to the valency. The atomic numbers of the rare earth and sulphur, and the crystal structure, are the only input data.

The central result of the present work is shown in Fig. 1. This is the calculated energy difference between the divalent and trivalent form of the rare-earth metals and their sulphides. A striking feature is the universality of these results: both elemental rare earths and their sulphides progress in a similar manner as a function of atomic number.

The calculated lattice sizes of the rare-earth metals and their sulphides, as a function of rare-earth atomic number, are shown in Fig. 2 and compared with experiment. These results represent a significant advance from the previous theory, which assumed a valence and calculated the size of the crystal^{3,4}. Here we calculate both the valence and the lattice size. An examination of the details of these calculations provides an understanding of valency in rare-earth materials.

The common feature of our calculations is an f band split into two sub-bands. One of these is occupied, and the other is unoccupied. The occupied f -electron states for both valence configurations form a band in a narrow energy range which lies well below the Fermi energy (E_F). The unoccupied band lies just above the Fermi energy. These unoccupied f bands hybridize with the s - d bands which straddle the Fermi energy. The result of this process is that the s - d bands acquire some f character, creating a different type of f electron which can participate in electronic bonding. Figure 3 shows schematic representations of the densities of states of the materials.

In the metals, the density of states consists of a broad band of s - d character crossing a narrow f band due to the f states that have not been corrected for self interaction. In the divalent configuration, the electron occupancy of two conduction electrons per atom is achieved with the Fermi level in the s - d band well below the f resonance. Consequently, no band-like f -electron states are occupied in this case. In contrast, in the trivalent configuration an occupancy of three conduction electrons per atom is required, which is achieved with the Fermi level at the f resonance. A number of $2+x$ electron states of s - d character are occupied together with $1-x$ f -like band electrons. As the f resonance has a substantial weight of one or more electrons, the Fermi level never falls above this resonance. Just how

many band-like f -electron states are occupied depends on the position of the f resonance relative to the s - d band. In the early rare-earth elements the f resonance is situated high up in energy, and the number $2+x$ is close to three, but as we proceed through the series the f resonance drops lower and x decreases. In samarium (Sm), where five f electrons are localized, it reaches its minimal value of 0.35. The next element, europium (Eu), prefers to fill its half shell in the f^7 configuration and transform to the divalent state. In the latter half of the rare-earth series starting with gadolinium (Gd), the same procedure occurs again with the spin-down f resonance starting high in energy and gradually dropping as we proceed to thulium (Tm), after which the next element, ytterbium (Yb), again prefers to complete its f shell in the divalent f^{14} configuration.

As seen in Fig. 3, the behaviour in the sulphides is similar. If the rare-earth ion is divalent there are two electrons to participate in bonding. Sulphur has four filled, and two empty, $3p$ bands, and the two electrons from the rare earth hybridize with the empty $3p$ bands, and fill them. The Fermi energy then falls in a gap between unfilled and filled bands, and the material is a semiconductor. As we proceed across the periodic table an extra electron is added to the localized f shell and the occupied f bands fall slightly in energy. However, the outer electronic structure of the s - d electrons is more or less unchanged. Therefore, bonding in the divalent sulphides is essentially unaffected by the addition of the f electrons.

In the trivalent rare-earth sulphides, there are three electrons available for bonding. Again, two of these hybridize with the sulphur $3p$ bands and fill them, but this still leaves one electron to be accommodated. This electron fills the first states above the sulphur $3p$ bands (Fig. 3). At the beginning of the rare-earth series, the empty $4f$ levels are held well above the Fermi energy. As the atomic number of the rare earth increases, the electron density associated with the filled $4f$ levels increases and they get closer to the nucleus. Therefore they are deeper in the nuclear potential well and their total energy is lowered. This also means that the ion looks more like a point charge, which enables the unfilled $4f$ levels to get closer to the Fermi energy. As this occurs, the nature of the outer electron changes. It is initially mostly s - d -like, but as the unoccupied f bands are pulled closer to the Fermi energy, hybridization between the itinerant s - d band and the more localized f band increases, and it gradually becomes energetically more favourable for the last electron to orbit the single ion, rather than be free to move through

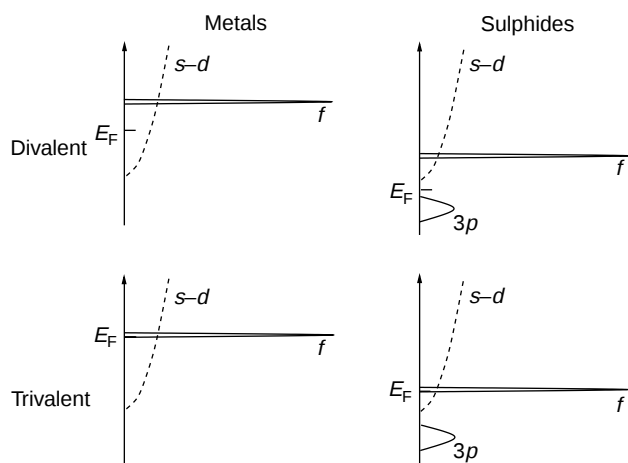


Figure 3 Schematic representation of the density of states around the Fermi energy for the divalent and trivalent rare-earth metals and their sulphides. This is the number of allowed electron states per unit energy range. At absolute zero, the states are filled up to the Fermi energy, E_F , and are empty above it.

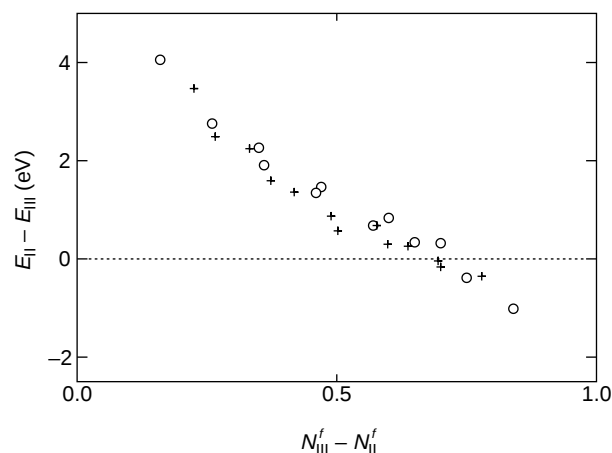


Figure 4 Energy difference (in eV) between divalent and trivalent states of rare earths and their sulphides versus the difference in their band-like f -electron number. Each point in this figure represents one rare-earth element (circles) or one rare-earth sulphide (crosses); the almost linear relationship confirms our discussions. The fact that the metals and sulphides fall on the same line strongly suggests that it is the number of delocalized f electrons on the rare earth that determines which valence is the more stable state. When the difference becomes greater than ~ 0.7 , the divalent state is the more stable.

the material. When we reach SmS, the state immediately above the Fermi energy contains more f character than s – d character, and the lowest-energy state has this electron participating in the occupied f subshell; that is, the divalent phase is favoured. This behaviour occurs in the light rare earths, from CeS to EuS, and is repeated in the heavy rare earths from GdS to YbS.

As we go from EuS to GdS in the trivalent state, there is a sudden change in the electronic structure. In trivalent EuS there is a sub-band of six occupied localized spin-up f states, plus one spin-up band state which is predominantly f -like and which is tied to the Fermi level. The unoccupied f bands are held well above the Fermi level by the magnetic splitting. When we go to the trivalent form of GdS, the seventh spin-up f band is localized and is no longer tied to the Fermi energy. The spin-up bands then fall closer to the nucleus, to the position they would occupy in the atom. The unoccupied f bands can then drop close to the Fermi energy.

There are two rather different types of f electron in these materials. The fully occupied f states are strongly localized and have the characteristics of core states. These determine the 'valence'. The outer electron is less well localized, meaning that the number of f electrons in these materials is not an integer. The difference in the total number of f electrons between the divalent and trivalent states changes with atomic number and compound. Hence the traditional view that the number of f electrons determines the valence is shown to be not well founded. Consequently, a valence transition is not, in general, a transition between two states with integer numbers of f electrons. Rather, it is a transition between two states with integer numbers of localized f electrons and an unspecified number of other f electrons.

Our discussion of valency implies that there is a direct relationship between the difference in the number of the less well localized f electrons and the difference in energy between the two valence states. This is plotted in Fig. 4: a linear relationship is indeed observed. □

Received 5 March; accepted 10 May 1999.

1. Jensen, J. & Mackintosh, A.R. *Rare Earth Magnetism* (Oxford Science Publications, Oxford, 1991).
2. Skriver, H. in *Systematics and Properties of the Lanthanides* (ed. Sinha, S.P.) 213–254 (Reidel, Dordrecht, 1983).
3. Johansson, B. Energy position of 4f levels in rare earth metals. *Phys. Rev. B* **20**, 1315–1327 (1979).
4. Delin, A., Fast, L., Johansson, B., Wills, J.M. & Eriksson, O. Method for calculating valence stability of lanthanide systems. *Phys. Rev. Lett.* **79**, 4637–4640 (1997).
5. Svane, A. & Gunnarsson, O. Transition-metal oxides in the self-interaction-corrected density-functional formalism. *Phys. Rev. Lett.* **65**, 1148–1151 (1990).
6. Temmerman, W.M., Svane, A., Szotek, Z. & Winter, H. in *Electronic Density Functional Theory: Recent Progress and New Directions* (eds Dobson, J.F., Vignale, G. & Das M.P.) 327–347 (Plenum, New York, 1998).
7. Jayaraman A. in *Handbook of the Physics and Chemistry of Rare Earths* Vol. 2 (eds Gschneider, K.A. Jr & Eyring, L.R.) 575–611 (North Holland, Amsterdam, 1979).
8. Syassen, K. Ionic monochalcogenides under pressure. *Physica B* **139 & 140**, 277–283 (1986).
9. Syassen, K., Winzen, H., Zimmer, H.G., Tups, H. & Leger, J. M. Optical response of YbS and YbO at high pressure and the pressure-volume relation for YbS. *Phys. Rev. B* **32**, 8246–8252 (1985).

Acknowledgements. We thank B. Johansson and O. Eriksson for discussions. This work was supported by the Psi-k European TMR network.

Correspondence and requests for materials should be addressed to P.S. (e-mail: P.Strange@phys.keele.ac.uk).

The electronic structure at the atomic scale of ultrathin gate oxides

D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt & G. Timp

Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

The narrowest feature on present-day integrated circuits is the gate oxide—the thin dielectric layer that forms the basis of field-effect device structures. Silicon dioxide is the dielectric of choice and, if present miniaturization trends continue, the projected oxide thickness by 2012 will be less than one nanometre, or about

five silicon atoms across¹. At least two of those five atoms will be at the silicon–oxide interfaces, and so will have very different electrical and optical properties from the desired bulk oxide, while constituting a significant fraction of the dielectric layer. Here we use electron-energy-loss spectroscopy in a scanning transmission electron microscope to measure the chemical composition and electronic structure, at the atomic scale, across gate oxides as thin as one nanometre. We are able to resolve the interfacial states that result from the spillover of the silicon conduction-band wavefunctions into the oxide. The spatial extent of these states places a fundamental limit of 0.7 nm (four silicon atoms across) on the thinnest usable silicon dioxide gate dielectric. And for present-day oxide growth techniques, interface roughness will raise this limit to 1.2 nm.

It is now technologically possible to produce metal oxide semiconductor field effect transistors (MOSFETs) with gates shorter than 50 nm and SiO₂ gate oxides less than 1.3 nm thick². Such a thin gate oxide is required to improve the drain-current response of the transistor to the applied gate voltage (allowing lower voltages to be used). As power dissipation at present limits the scale of integration, lowering the power supply voltage becomes the key to increasing integration and improving integrated-circuit performance. The performance of the gate oxide therefore becomes the limiting factor when manufacturing very-large-scale integrated circuits. As a practical alternative to SiO₂ (or its nitrogenated derivatives), providing a higher dielectric constant or a reduced leakage current, has not been identified yet¹, it is crucial to the future of very-large-scale integration (VLSI) to discover the practical limits on the thickness of the SiO₂ gate oxide.

There are two fundamental considerations. First, the roughness of the interface must be controlled at an atomic scale if such thin oxides are to prove practical. The leakage current through a 1-nm-thick oxide increases by about a factor of 10 for every 0.1-nm increase in the root-mean-square roughness. This leakage current, in conjunction with the subthreshold leakage, is the most important figure of merit in a MOSFET. Second, a single layer of silicon and oxygen has the incorrect topology to reproduce the local electronic structure of bulk silicon dioxide³. The question is then how thick must a silicon dioxide layer be before its bulk electrical properties can be obtained? The presence of an intrinsic transition region (which may be a substoichiometric oxide—the 'suboxide') between bulk Si and the bulk-like SiO₂ will place a fundamental limit on drive current by limiting the minimum thickness. Attempts to measure the width of the transition region have given answers that range from structurally abrupt (for molecular beam epitaxy on an atomically flat substrate⁴) to a chemical thickness of 0.3–0.5 nm (for thermally grown⁵ oxides). A comprehensive review of earlier work in the field (and the consequences for electronic properties of the interface) is given in ref. 6. However, there is considerable disagreement as to the precise structure and chemical composition of this suboxide^{7,8}.

Even if the interface structure were known, the connection between the physical arrangement of atoms at the interface and their electrical properties is neither direct nor obvious. Here we focus on measuring the electronic states that directly determine the electrical properties of the interface, which we do with atomic-scale electron-energy-loss spectroscopy (EELS)^{9–12}. We use EELS to map the unoccupied electronic density of states by site, atom-column and atomic species. These measurements give localized information about both chemical composition and electronic properties. The work of Batson¹¹ is particularly relevant for the present study as he demonstrated that a usable Si L edge EELS signal can be obtained from an atomic-sized probe at a Si/SiO₂ interface. To improve the contrast and sensitivity, we found it necessary to use the higher-energy (but weaker) oxygen K edge which is more localized than the Si L edge¹³.

The O K and Si L_{2,3} EELS edges provide information on the

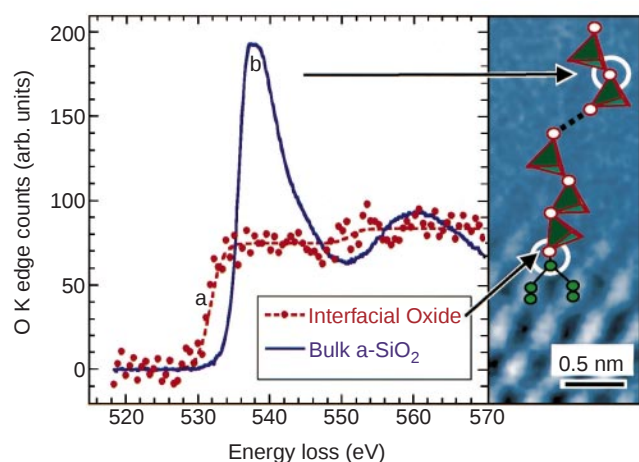


Figure 1 The measured EELS O K edges of bulk a-SiO₂ and for O atoms at an atomically smooth interface between [100]Si and native a-SiO₂. The 3-eV reduction in the edge onset at the interface aligns the unoccupied O interfacial states with the Si conduction band edge, as would be expected for induced gap states. This pre-peak (a) can be thought of as tunnelling states leaking in from the bulk Si. The reduction in the first bulk peak (b) is discussed in the text. The figure to the right is the annular dark field (ADF) image of the interface between the a-SiO₂ and the silicon substrate. The image is a projection of the ~100 atoms in the path of the beam, so no structural information can be obtained from the amorphous region where the atoms are not aligned along the beam direction. At a rough interface, the bulk lattice would gradually fade into the background over many lattice spacings. The measurements were performed on the Cornell VG-HB501 100 kV STEM equipped with a cold field-emission gun and a McMullan-style parallel EELS spectrometer. The microscope has been modified to achieve high energy-drift stability (<0.2 eV min⁻¹) and spatial-drift stability (<0.05 nm min⁻¹).

unoccupied O-*p*, and Si-*s,d* electronic densities of states (DOS) respectively¹⁴. The effect of the 2*p* core hole on the Si L edge is significant, producing a strong exciton. However, the 1*s* core hole on the O K edge does not introduce any new features at the 1-eV energy resolution used in this study (as we verified by comparison of *ab initio* calculations of the ground-state DOS to the EELS spectra (J. Neaton and D.A.M., unpublished results)). This allows a single-particle interpretation of the EELS spectra, which are proportional to local densities of states partitioned in three ways: by site (as the incident probe is localized); chemical species (as each element has unique core level binding energies); and angular momentum (from the dipole selection rules)^{15,16}. In a scanning transmission electron microscope (STEM), the EELS measurements are made at internal interfaces, not free surfaces, by passing the 100-keV electron beam (0.2–0.5 nm diameter) through a thin film^{9–12,17}. The film is chosen to be thick enough that surface states produce a negligible fraction of the transmitted signal, but also thin enough that multiple scattering is not significant. The interface is oriented parallel to the beam, so that a column of atoms in the interface plane can be measured separately from any atoms in adjacent columns. But because the interface is viewed in projection, any interfacial roughness (especially on length scales thinner than the sample) can lead to an apparent broadening of the interface. We therefore also use X-ray reflectivity to obtain independent measurements of the interface roughness.

We found that the thermal processing needed to grow the gate oxides roughened the interfaces (possibly by microfaceting as the step density is not altered). Consequently, we also searched for a model system containing atomically abrupt interfaces, before examining the thermal oxides. Native oxides which were formed in a dry atmosphere on epitaxial silicon layers produced by chemical vapour deposition (CVD) on [001] silicon could be less than 0.8 nm

thick and have very smooth Si/SiO₂ interfaces. The native oxides were prepared for microscopy by adding a protective overlayer of amorphous silicon (a-Si), resulting in Si/SiO₂/a-Si sandwich. The O K edge recorded at these smooth interfaces is strikingly different from that for bulk SiO₂ (Fig. 1). First, the edge onset (a in Fig. 1) is reduced by 3 eV at the interface with respect to the bulk. As the O K edge reflects the portion of the conduction band projected onto the probed O atom, the reduced edge onset implies a reduced bandgap (which will probably increase the local dielectric constant and electrical conductivity⁶). Second, the sharp peak (b in Fig. 1), which is the first extended-fine-structure peak in the bulk near-edge spectrum, is absent at the interface. X-ray absorption studies of the O K edge in quartz, cristobalite and coesite (three different forms of SiO₂ which have the same nearest-neighbour topologies and O–Si–O bond angles but differ mainly in the Si–O–Si bond angles and dihedral angles) all show the same near-edge structure¹⁸. The same edge shape is also present in GeO₂ for the quartz structure, but not for rutile, suggesting that it is the position of the O atoms and not the cations that determines the shape of peak b. Generally, the sharp peak b is identified as arising from O–O scattering and it decreases in intensity as the number of O second-nearest neighbours around the excited O atom are reduced¹⁹. Essentially, silicon is a much weaker scatterer than O and does not produce strong extended-fine-structure oscillations. We expect a reduction in this peak's intensity even for an atomically abrupt interface as the last layer of O atoms would always lose half its O second-nearest neighbours. (A similar effect can be seen at Cu/MgO interfaces¹⁷). However, the almost complete absence of this peak implies that more O second-nearest neighbours are missing—that is, the last atom is in a silicon-rich environment. We will use this feature later to estimate the width of the suboxide in the thermally grown gate oxides.

A similar interfacial spectrum is observed at the interface between Si and the thermally grown oxides. Figure 2 shows EELS spectra recorded point by point across a gate stack whose oxide thickness was nominally measured at 1 nm by ellipsometry. The spectrum from each point was decomposed into a linear combination of the bulk and interface signals from Fig. 1, and this decomposition was used to identify the localized relative fractions of bulk-like and interfacial oxide signals (plotted in Fig. 3a). We note that 60 ± 6% of the total oxygen signal is generated by the interfacial atoms (whose local electronic structure is very different from the bulk). From X-ray reflectivity measurements, we find that the substrate/oxide interface roughness has a standard deviation of $\sigma_r = 0.1$ nm over length scales from 0.1 to 1,000 nm. The projected peak–peak roughness (that is, the distance from minimum to maximum excursion of the silicon substrate) expected in the EELS measurements is then $6\sigma_r = 0.6$ nm. Such roughness will spread out the interfacial signal, such as that in Fig. 3a, without changing its total area. Consequently, the apparent broadness of the lower interface is dominated by the expected 0.6-nm interface roughness, plus the chemical width of the interface (and the 0.27-nm-wide probe added in quadrature). The width of the upper interface between the oxide and the a-Si is half that of the lower interface, but the interfacial areas are the same. This suggests that the main difference between the thermally grown lower interface and the upper interface produced by CVD of a-Si is that the upper interface is much smoother.

The gate stack was then annealed at 1,050 °C for 10 s—a thermal budget typical of that used in device processing. Figure 3b shows that after annealing, the upper interface (which has now been converted to polycrystalline silicon) becomes as rough as the lower. However, the fraction of interfacial signal is essentially unchanged at 50 ± 5%. The most significant difference is that the two interfacial regions now overlap. As the interfacial signal (feature a of Fig. 1) is associated with tunnelling states (discussed below), a large leakage current is expected and observed. This leakage current

of 10 A cm^{-2} is higher than desirable for use in integrated circuits¹, but the device is still able to function as a transistor², producing drive currents in excess of $1 \text{ mA per } \mu\text{m}$. By increasing the gate oxide thickness to a point where the two interfacial regions no longer overlap, the tunnelling can be reduced. Figure 3c shows a thicker gate oxide (1.8 nm ellipsometric thickness) where the measured electrical-leakage current has been suppressed by 6–7 orders of magnitude. The width of the interface region is not increased in the thicker oxide.

We now estimate the width of the suboxide (that is, the region of substoichiometric oxide) from the line profile of the interface states by recalling that the extended fine structure on the O K edge (especially peak b) is sensitive to the number of second-nearest neighbours that are O atoms¹⁹. We need to correct for both the interfacial roughness and the fact that the EELS signal is sensitive to second-nearest neighbours as well as nearest neighbours (unlike X-ray photoelectron spectroscopy, XPS). The roughness (and finite instrument resolution) can be accounted for by using the full-width at half-maximum (FWHM) of the total oxygen line-scan (Fig. 3) as a measure of the spatial width of the oxide.

The FWHM of the line profile is independent of interfacial roughness and instrumental spatial resolution, provided that the roughness profile is symmetric about the centre of each interface (as is the case in Fig. 3). For Fig. 3b, the FWHM of the bulk-like oxide is $0.85 \pm 0.05 \text{ nm}$. From the ratio of areas of interface and bulk-like

oxide, each interfacial region (defined as that region lacking peak b on the O K edge and hence having fewer oxygen second-nearest neighbours) has a FWHM of $0.43 \pm 0.05 \text{ nm}$. The interface thus defined can never be sharper than 0.27 nm (as this is the distance between O second-nearest neighbours, and the last O atom in contact with the bulk Si will always be missing some second-nearest neighbours even at a perfect interface). Consequently, the suboxide cannot be thicker than $0.43 - 0.27 = 0.16 \text{ nm}$ (that is, 1–2 monolayers thick). This upper limit placed by the EELS measurements is consistent with the suboxide thickness inferred indirectly from XPS⁵.

The additional electronic states at the interface (feature a in Fig. 1) which appear at energies below the bulk SiO_2 conduction band edge are roughly aligned with the bulk silicon conduction band. We identify these as induced gap states²⁰ resulting from the exponential decay of the silicon conduction band wavefunctions into the oxide (consequently there will be no states in the bandgap of bulk silicon). These evanescent states should be a general feature of any Si/ SiO_2 interface and should not be very sensitive to the detailed atomic structure. The additional states in the SiO_2 bandgap near the interface also imply an altered dielectric constant there. The

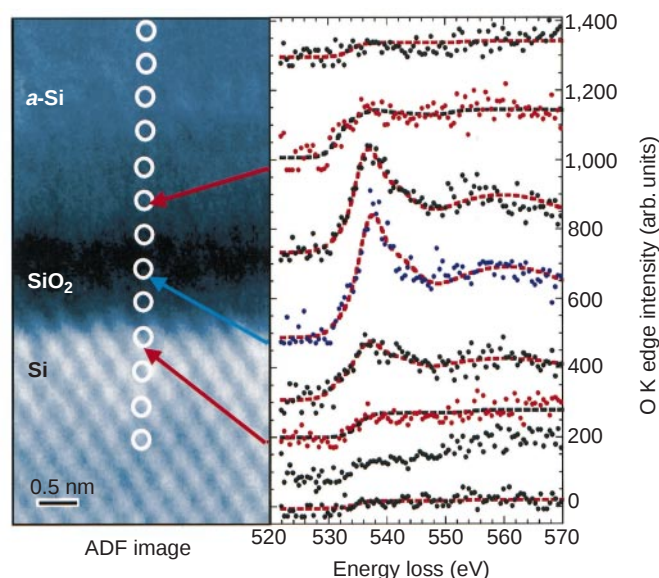


Figure 2 EELS spectra recorded point by point across a gate stack containing a thin gate oxide. The annular dark field (ADF) image (left panel) shows where each spectrum was taken. The [100] silicon substrate is in the lower half, the gate oxide in the middle and the deposited a-Si layer is in top half. The right panel shows the background-corrected O K edges. The smooth curves are the best fits using a mixture of the bulk and interfacial spectra from Fig. 1. The spectra were recorded with a 4-s exposure per point to ensure that the dose was less than the threshold for detectable radiation damage. In growing the gate stacks, particular care was taken to select (using scanning tunnelling microscopy) silicon wafers with exceptionally smooth surfaces before oxidation. The [001]-oriented wafers were cleaned for 15 s in a 15:1 HF solution followed by a 30 s UV-CL clean. The gate oxide was produced by rapid thermal oxidation at $1,000^\circ\text{C}$. Chemical vapour deposition (CVD) of a-Si at 550°C on top of the gate oxide was used as the base of the gate stack, which was subsequently annealed and implanted with dopants. For microscopy, cross-sectioned specimens were dimpled and then ion-milled to electron transparency. The EELS spectra were recorded simultaneously with the ADF signal, under conditions optimized for atomic-resolution imaging^{9–12,17}. For a 10-nm-thick sample, the ADF resolution is 0.23 nm and the EELS spatial resolution at the O K edge is reduced to 0.26 nm due to delocalization¹⁹.

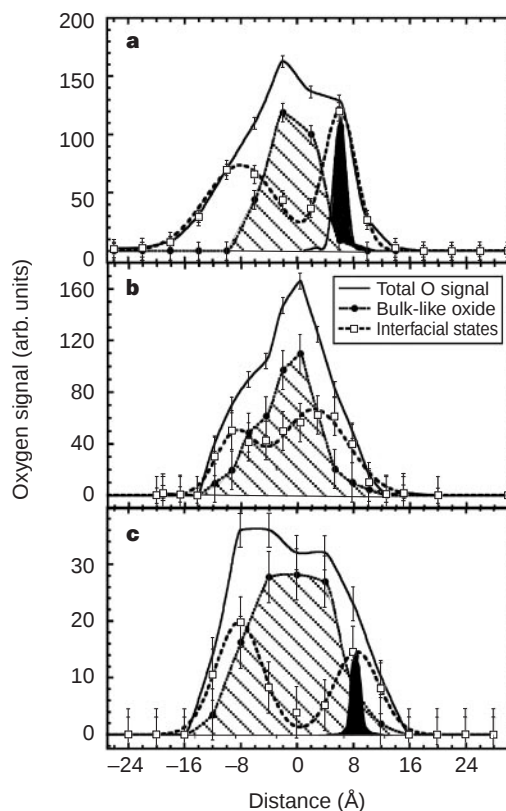


Figure 3 Oxygen bonding profiles from O K edge EELS. The substrate is on the left, and the gate contact is on the right: **a**, for a c-Si/ SiO_2 /a-Si stack before annealing (c-Si indicates crystalline Si). The nominal ellipsometric thickness measured after oxidation was 1 nm . From EELS, the total O signal has a FWHM of 1.6 nm and the bulk-like oxide (hatched area) has a FWHM of 0.85 nm . **b**, After annealing for 10 s at $1,050^\circ\text{C}$, the a-Si is converted to polycrystalline silicon, roughening the upper interface. The total O signal has a FWHM of 1.3 nm and the bulk-like oxide has a FWHM of 0.85 nm . Although the amount of interfacial oxide has not increased, the increased roughness has resulted in an overlap of the two interfacial regions, almost shorting-out the gate oxide (the leakage current is 10 A cm^{-2}). **c**, In a slightly thicker gate oxide (1.8 nm ellipsometric thickness), the two interfacial signals no longer overlap and the leakage current is reduced to $10^{-5} \text{ A cm}^{-2}$. The total O signal has a FWHM of 2.1 nm and the bulk-like oxide has a FWHM of 1.6 nm . The inelastic point-spread function, which gives a measure of the probe resolution, is shown as black in **a** and **c**.

interfacial dielectric constant will probably be between that of Si and SiO₂. The altered dielectric layers are not accounted for in our ellipsometric measurements, in which we assume that the oxide has only the dielectric constant of bulk SiO₂. This is probably why ellipsometry has underestimated the width of the oxides in Fig. 3.

The probe localization is obtained by constructing a wavepacket with a transverse momentum spread of more than a reciprocal lattice vector, and consequently all electronic momentum information is lost (as required by the uncertainty principle). Therefore these evanescent states responsible for tunnelling through the oxide and the states from the extended conduction band are treated on an equal footing, and cannot be separated in such a local measurement. In the simplest model, the silicon wavefunctions decay exponentially into the oxide barrier with a decay length for the evanescent states, $\lambda(E)$, determined by the energy difference between the interfacial state (E) and the conduction band edge of bulk SiO₂, (E_c), as $\lambda(E) = \hbar/\sqrt{E_c - E}$. The tunnelling current depends on the overlap of the evanescent states from either interface. A satisfactory tunnelling barrier is formed when the oxide thickness t is 6λ . This sets an absolute minimum thickness of $t_{\min} = 0.7$ nm for an ideal SiO₂ gate oxide. Interfacial roughness adds another $6\sigma_r$ to $t_{\min}$. The smallest roughness for our thermally grown oxides was $6\sigma_r = 0.6$ nm which puts a lower limit of 1.2 nm on the practical SiO₂ gate oxide thickness. The induced gap states also place severe constraints on the minimum allowed thickness for alternative dielectrics, many of which have large dielectric constants, but reduced bandgaps and hence longer decay lengths. Furthermore, there is the possibility of a reaction between the dielectric and the silicon substrate to form a silicon oxide interlayer. If the interlayer thickness exceeds 1.3 nm (and a typical native oxide is 2 nm thick), the gate capacitance is less than what could be obtained with a pure SiO₂ gate oxide. □

Received 20 January; accepted 8 April 1999.

1. Semiconductor Industry Association *The National Technology Roadmap for Semiconductors* 71–81 (Sematech, Austin, 1997).
2. Timp, G. et al. in *IEDM Technical Digest* 615–618 (IEDM, San Francisco, 1998).
3. Cryot-Lackmann, F. Sur le calcul de la cohésion et de la tension superficielle des métaux de transition par une méthode de liaisons fortes. *J. Phys. Chem. Solids* **29**, 1235–1243 (1968).
4. Ourmazd, A., Taylor, D. W., Rentschler, J. A. & Bevk, J. Si to SiO₂ transformation: interfacial structure and mechanism. *Phys. Rev. Lett.* **59**, 213–216 (1987).
5. Himpfel, F., McFeely, F. R., Taleb-Ibrahimi, A., Yarmoff, J. A. & Hollinger, G. Microscopic structure of the SiO₂/Si interface. *Phys. Rev. B* **38**, 6084–6096 (1988).
6. Grunthaner, F. J. & Grunthaner, P. J. Chemical and electronic structure of the Si/SiO₂ interface. *Mater. Sci. Rep.* **1**, 65–160 (1986).
7. Pasquarello, A., Hybertsen, M. S. & Car, R. Theory of Si 2p core-level shifts at the Si(001)-SiO₂ interface. *Phys. Rev. B* **53**, 10942–10950 (1996).
8. McFeely, F. R., Zhang, K. Z., Banaszak Holl, M. M., Lee, S. & Bender, J. E. An inquiry concerning the principles of the Si 2p core-level photoemission shift assignments at the Si/SiO₂ interface. *J. Vac. Sci. Technol. B* **14**, 2824–2830 (1996).
9. Muller, D. A., Tzou, Y., Raj, R. & Silcox, J. Mapping sp^2 and sp^3 states of carbon at sub-nanometre spatial resolution. *Nature* **366**, 725–727 (1993).
10. Muller, D. A., Subramanian, S., Sass, S. L., Silcox, J. & Batson, P. E. Near atomic scale studies of electronic structure at grain boundaries in Ni₃Al. *Phys. Rev. Lett.* **75**, 4744–4747 (1995).
11. Batson, P. E. Simultaneous STEM imaging and electron energy-loss spectroscopy with atomic column sensitivity. *Nature* **366**, 727–728 (1993).
12. Browning, N. D., Chisholm, M. M. & Pennycook, S. J. Atomic-resolution chemical analysis using a scanning transmission electron microscope. *Nature* **366**, 143–146 (1993).
13. Muller, D. A. & Silcox, J. Delocalization in inelastic scattering. *Ultramicroscopy* **59**, 195–213 (1995).
14. Egerton, R. F. *Electron Energy Loss Spectroscopy in the Electron Microscope* 2nd edn (Plenum, New York, 1996).
15. Colliex, C. & Jouffrey, B. Diffusion inelasticité des électrons dans une solide par excitation de niveaux atomiques profonds. *Phil. Mag.* **25**, 491–514 (1972).
16. Müller, J. E. & Wilkins, J. Band-structure approach to the x-ray spectra of metals. *Phys. Rev. B* **29**, 4331–4348 (1984).
17. Muller, D. A. et al. Atomic scale observations of metal-induced gap states at [222] MgO/Cu interfaces. *Phys. Rev. Lett.* **80**, 4741–4744 (1998).
18. Brown, G. E. Jr, Waychunas, G. A., Stohr, J. & Settle, F. Near-edge structure of oxygen in inorganic oxides: effect of local geometry and cation type. *J. Phys. C*, (Colloque C8) 685–689 (1986).
19. Wallis, D., Gaskell, P. H. & Brydson, R. Oxygen K near-edge spectra of amorphous silicon suboxides. *J. Microsc.* **180**, 307–312 (1993).
20. Zangwill, A. *Physics at Surfaces* (Cambridge Univ. Press, New York, 1988).

Acknowledgements. We thank D. R. Hammann, M. S. Hybertsen, P. Rez, J. Neaton and B. Batlogg for discussions, and J. Silcox and M. Thomas for access to the Cornell Center for Materials Research STEM. Funding for the operation and acquisition of the STEM was provided by the NSF Upgrades were funded by the US Air Force Office of Scientific Research. The X-ray diffraction was performed on X16B at the National Synchrotron Light Source.

Correspondence and requests for materials should be addressed to D.A.M. (e-mail: davidm@bell-labs.com).

Molecular mechanistic origin of the toughness of natural adhesives, fibres and composites

Bettye L. Smith^{*,†}, Tilman E. Schäffer^{†,‡}, Mario Viani^{*}, James B. Thompson^{*}, Neil A. Frederick^{*}, Johannes Kindt^{*}, Angela Belcher[§], Galen D. Stucky^{||}, Daniel E. Morse[†] & Paul K. Hansma^{*}

^{*} Department of Physics, ^{||} Department of Chemistry and Materials, and

[†] Department of Molecular, Cellular, and Developmental Biology,

University of California at Santa Barbara, California 93106, USA

[‡] Department of Molecular Biology, Max-Planck-Institute for Biophysical

Chemistry, 37070 Göttingen, Germany

[§] Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, USA

[†] These authors contributed equally to this work

Natural materials are renowned for their strength and toughness^{1–5}. Spider dragline silk has a breakage energy per unit weight two orders of magnitude greater than high tensile steel^{1,6}, and is representative of many other strong natural fibres^{3,7,8}. The abalone shell, a composite of calcium carbonate plates sandwiched between organic material, is 3,000 times more fracture resistant than a single crystal of the pure mineral^{4,5}. The organic component, comprising just a few per cent of the composite by weight⁹, is thought to hold the key to nacre's fracture toughness^{10,11}. Ceramics laminated with organic material are more fracture resistant than non-laminated ceramics^{11,12}, but synthetic materials made of interlocking ceramic tablets bound by a few weight per cent of ordinary adhesives do not have a toughness comparable to nacre¹³. We believe that the key to nacre's fracture resistance resides in the polymer adhesive, and here we reveal the properties of this adhesive by using the atomic force microscope¹⁴ to stretch the organic molecules exposed on the surface of freshly cleaved nacre. The adhesive fibres elongate in a stepwise manner as folded domains or loops are pulled open. The elongation events occur for forces of a few hundred piconewtons, which are smaller than the forces of over a nanonewton required to break the polymer backbone in the threads. We suggest that this 'modular' elongation mechanism might prove to be quite general for conveying toughness to natural fibres and adhesives, and we predict that it might be found also in dragline silk.

We have looked for the mechanism behind the toughness of the organic adhesives and fibres, and, in particular, at the nacre in abalone shells. Analysis of the insoluble organic matrix from the abalone shell revealed a fibrous core in the interlamellar sheets placed between successive nacre tablets^{5,15,16}, which probably serve as an adhesive holding the tablets together. The organic adhesive is readily apparent when the tablets are pulled apart (Fig. 1). At least one protein, lustrin A, has been isolated from this insoluble organic matrix. The complementary DNA and translated protein sequence reveal that the structure of this protein consists of about 10 alternating and highly conserved cysteine-rich and proline-rich domains, demonstrating that the structure is highly modular¹⁷. Immunohistochemical analysis of the fibres (Fig. 1) revealed lustrin A to be a component of the adhesive between the nacre mineral tablets.

Rief et al.^{18,19} demonstrated that the modular structure of a single molecule can be examined by attaching the molecule between a flat surface and a cantilever of an atomic force microscope (AFM). By pulling on the protein titin and plotting the force versus extension curves, these authors measured the force required to break open the individual subunit in its modular structure. As the titin

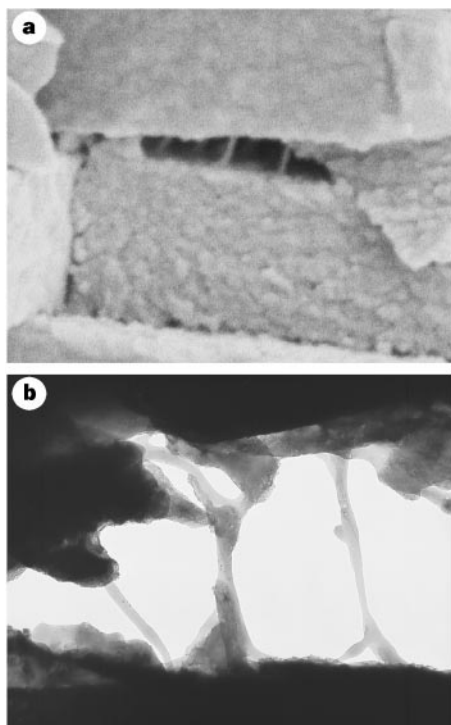


Figure 1 Scanning and transmission electron micrographs of a freshly cleaved abalone shell, showing adhesive ligaments formed between nacre tablets. **a**, Scanning electron micrograph of a freshly cleaved abalone shell showing adhesive ligaments formed between consecutive abalone nacre tablets on exertion of mechanical stress. The tablets are ~ 400 nm thick. **b**, Transmission electron micrograph of another cleaved abalone shell, showing the adhesive ligaments between nacre tablets. The space between the tablets is ~ 600 nm. Thus the ligaments can lengthen to many times the original spacing between the tablets, which is of the order of 30 nm.

molecule was stretched, the force–extension curve revealed a sawtooth pattern. Using titin constructs, Rief *et al.*¹⁹ demonstrated that every peak in the sawtooth pattern corresponded to a single domain unfolding. Apparently, it is the cumulative effect of the intermediate-strength hydrophobic bonds in the immunoglobulin-like domains of titin which contribute to the sawtooth force–distance curves found for this molecule^{20,21}. Studies using optical tweezers corroborated these results^{22–24}. These initial studies demonstrate that individual titin subunits unfold one at a time, suggesting that this technique affords an opportunity to delineate some of nature's mechanisms in building modular elastic fibres.

We have used this technique to investigate the molecules that hold abalone shells together. Figure 2 shows force–extension curves for the organic material exposed on a freshly cleaved nacre surface. Breaking forces of the order of 100–400 pN are seen in these curves. The hysteresis observed after a complete pulling cycle demonstrates that work has been done on the shell. This work is irreversibly dissipated as heat, and the area between the retracting and the approaching parts of the curve quantifies this heat. In this case, the dissipated heat is of the order of $(0.4\text{--}1) \times 10^{-17}$ J per cycle (Fig. 2). The sawtooth-like force–extension curves, and the observation that we could repeatedly unfold molecule(s) without touching the surface between pulls, suggest that bonds of some kind are breaking and reforming. As lustrin A is present on that surface, it is possible that we observed unfolding of this protein. However, and this is the key point, the mechanism behind the strength and toughness of the adhesive is revealed by the force–extension curves and does not depend on the identification of the specific molecules involved. This behaviour may reflect the successive opening of intra-chain loops or folded domains within a single molecule, or the successive release

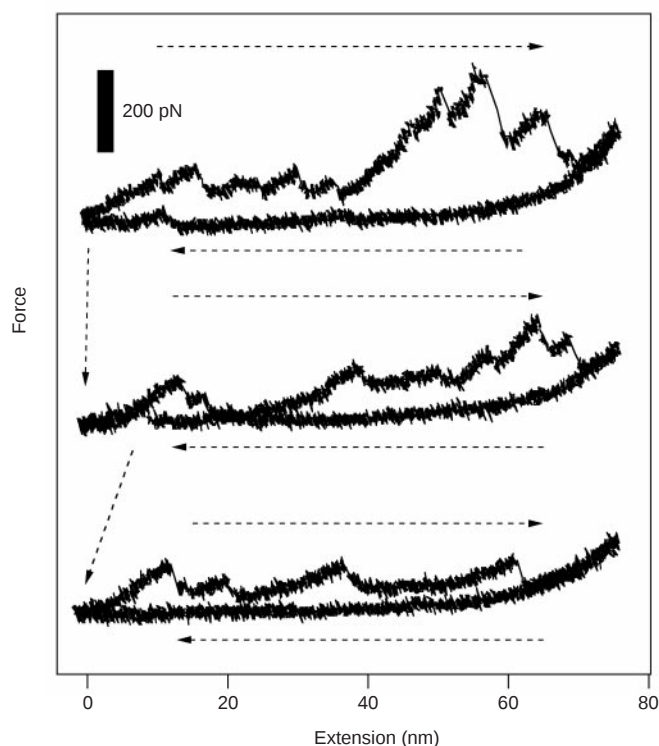


Figure 2 Consecutive force–extension curves, obtained using an atomic force microscope, from pulling on a freshly cleaved abalone nacre surface. Rupture events, with a sawtooth appearance, are visible in each of the curves. The surface was not touched between pulls, strong evidence that some refolding took place, possibly of domains in lustrin A. The approach and retract curves show hysteresis, indicating that the rupture events dissipate energy.

of sacrificial inter-chain bonds holding a crosslinked multichain matrix together.

To demonstrate the effect that a material with this type of force–extension curve can have on the properties of an adhesive made from modular fibres, we consider three different cases. Gluing materials together with conventional adhesives has traditionally involved either relatively stiff adhesives such as epoxies, or elastic adhesives such as silicon adhesives. When pulling on two surfaces glued together with a short molecule, the pulling force increases rapidly with only a little extension of the molecule (Fig. 3). First, a perfect stiff adhesive would be a short molecule bound to each surface by strong (that is, covalent or ionic) bonds and the molecules of the adhesive itself would be held together with strong bonds. Thus the break strength of each adhesive molecule would be the force required to break a strong bond: of the order of one nanonewton (estimated by dividing one electron volt by an extension of one ångström). For a material with many strongly bound molecules in parallel, the macroscopic tensile strength is expected to be of the order of several gigapascals. This is the order of magnitude for the breaking force of strong polymers such as Kevlar^{25,26}. The fracture toughness of such materials is rather small, however, even though the forces are large. This can be understood by considering the area under the force–extension curve (Fig. 3), which is the energy required to break the material. Because those stiff materials have a small elastic strain, the extension over which the force must be exerted until it breaks is small. Therefore the area under the curve, or the energy required to break the material, is small (Fig. 3).

Second, in contrast to this behaviour, the idealized curve for an elastic fibre made of long molecules shows that the force increases

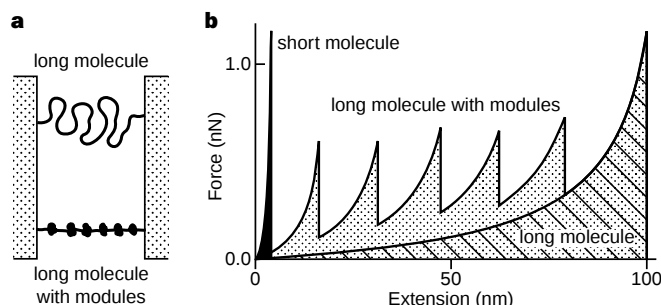


Figure 3 Model of long polymers and force-extension curves for different kinds of polymers. **a**, Diagram of long polymers behaving as entropic springs. The lower molecule is compacted, with many domains that are held together with intermediate-strength sacrificial bonds. **b**, Force-extension curves for three different kinds of molecules. A short molecule (solid curve) resists pulling up to a high force before it breaks at small extensions. The energy required to break this short molecule (area under the curve) is small. A long molecule (hatched curve) behaves like an entropic spring and yields to the pulling force up to large extensions. The energy to break the long molecule is larger than that for the small molecule, but the forces at low extensions are small. But a long molecule that is compacted into domains (the stippled plus the hatched curve) that are held together with intermediate-strength bonds resists pulling already at small extensions. Before the molecule's backbone can break at forces of the order of 1 nN, modules unfold at intermediate forces (0.1–0.7 nN). Repeated unfolding of molecules allows stretching up to large extensions, required significant energy. Therefore, a long molecule that is compacted into domains that are held together with intermediate-strength bonds combines both high (tensile) strength and high toughness.

slowly as the elastic material is stretched to the point at which the elastic limit is reached (Fig. 3)^{27,28}. Then the force increases rapidly for further extension until it breaks. This break will again occur at a force of the order of one nanonewton per molecular chain, assuming each chain is bound to each surface with a strong bond and is itself held together by strong bonds. Contrary to the case of short, inelastic molecules, the pulling force must be applied over much larger extensions. Therefore, the area under the force-extension curve would be larger (Fig. 3), and thus more energy would be needed to break the material. Unfortunately, the technology does not at present exist for making such an idealized elastic material. Real elastic materials such as rubbers have tensile strengths that correspond to breaking forces per molecule of the order of 0.1% of the theoretical maximum. But even if ideal elastic materials could be produced, there would still be a significant advantage to incorporating nature's mechanism of tying up most of the molecule's length with intermediate-strength sacrificial bonds.

Third, an adhesive or fibre that combines the best of both worlds would display a force-extension curve that rises to a large force quickly, but then maintains that force over large extensions. It would consist of long molecules with many domains that are either folded or looped, with intermediate-strength bonds (of the order of 0.1–0.7 nN) compacting the molecule. These intermediate-strength bonds are stronger than individual hydrogen bonds or van der Waals bonds, but weaker than covalent bonds. Alternatively, these may represent the cumulative effect of multiple weak bonds (either intra- or interchain) acting in concert. For such a molecule, the force rises quickly with extension (Fig. 3). But then, when the force rises to a significant fraction of the force required to break a strong bond and threatens to break the backbone of the molecule, a domain unfolds or a loop is opened, therefore avoiding the breaking of a strong bond in the backbone.

On further pulling, the force then again rises to a significant fraction of the force required to break a strong bond, but again, before a strong bond is broken, another domain unfolds. The process continues until all domains are unfolded (or loops are opened) and finally a strong bond breaks. The net result is to sustain

a large force over the pulling extension, making it 'strong', while producing a relatively large area under the force-extension curve, making it 'tough' as well.

An approximate analogy to this process is found in Greek mythology: Sisyphus was condemned by Zeus to push a heavy rock up a mountain only to have it slip out of his hands and roll back to the bottom just before he reached the summit. In such a fall, the rock would dissipate its potential energy into heat, and Sisyphus had to start all over again. The case of extending a modular fibre is analogous. One needs to pull hard, and do work, but before the breaking point (the 'summit') is reached, a domain unfolds or a loop opens, and the energy stored in the fibre is dissipated as heat. Then, the fibre has to be pulled on again, until the next domain breaks and so on, until no folded domains are left and the fibre at least breaks. (We note that if the force is relaxed before the fibre breaks, the domains or loops can reform as shown in Fig. 2.) In natural materials, proteins provide the modular fibre: but this mechanism should work for other polymers, as long as it is possible to compact the length of the polymers using intermediate-strength bonds that break sacrificially with forces comparable to, but less than, the force of a nanonewton or so that breaks strong bonds in the polymer backbone. □

Received 26 February; accepted 20 April 1999.

- Hinman, M., Dong, Z., Xu, M. & Lewis, R. V. in *Biomolecular Materials* (eds Viney, C., Case, S. T. & Waite, J. H.) 25–34 (Materials Research Soc., Pittsburgh, 1993).
- Qin, X., Coyne, K. J. & Waite, J. H. A novel natural copolymer: A collagenous molecule from mussel byssus contains silk fibroin-like domains. *Am. Zool.* **37**, 125A (1997).
- Waite, J. H., Qin, X. X. & Coyne, K. J. The peculiar collagens of mussel byssus. *Matrix Biol.* **17**, 93–106 (1998).
- Currey, J. D. Mechanical properties of mother of pearl in tension. *Proc. R. Soc. Lond. B* **196**, 443–463 (1977).
- Jackson, A. P., Vincent, J. F. & Turner, R. M. The mechanical design of nacre. *Proc. R. Soc. Lond. B* **234**, 415–440 (1988).
- Heslot, H. Artificial fibrous proteins: A review. *Biochimie (Paris)* **80**, 19–31 (1998).
- Vollrath, F. & Edmonds, D. T. Modulation of the mechanical properties of spider silk by coating with water. *Nature* **340**, 305–307 (1989).
- Qin, X. X., Coyne, K. J. & Waite, J. H. Tough tendons. Mussel byssus has collagen with silk-like domains. *J. Biol. Chem.* **272**, 32623–32627 (1997).
- Watabe, N. & Wilbur, K. M. (eds.) *The Mechanisms of Biomineralization in Invertebrates and Plants* (Univ. South Carolina Press, Columbia, SC, 1976).
- Weiner, S. Organization of extracellularly mineralized tissues: a comparative study of biological crystal growth. *CRC Crit. Rev. Biochem.* **20**, 365–408 (1986).
- Jackson, A. P., Vincent, J. F. V. & Turner, R. M. A physical model of nacre. *Composites Sci. Technol.* **36**, 255–266 (1989).
- Clegg, W. J., Kendall, K., Alford, N. M., Button, T. W. & Birchall, J. D. A simple way to make tough ceramics. *Nature* **347**, 455–457 (1990).
- Almqvist, N. et al. Methods for fabricating and characterizing a new generation of biomimetic materials. *Mater. Sci. Eng. C* **7**(1), 37–43 (1999).
- Binnig, G., Quate, C. F. & Gerber, C. The atomic force microscope. *Phys. Rev. Lett.* **56**, 930–933 (1986).
- Addadi, L. & Weiner, S. Biomineralization—A pavement of pearl. *Nature* **389**, 912–915 (1997).
- Schäffer, T. E. et al. Does abalone nacre form by heteroepitaxial nucleation or by growth through mineral bridges? *Chem. Mater.* **9**, 1731–1740 (1997).
- Shen, X., Belcher, A. M., Hansma, P. K., Stucky, G. D. & Morse, D. E. Molecular cloning and characterization of lustrin A, a matrix protein from shell and pearl nacre of *Haliotis rufescens*. *J. Biol. Chem.* **272**, 32472–32481 (1997).
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. Single molecule force spectroscopy on polysaccharides by atomic force microscopy. *Science* **276**, 1295–1297 (1997).
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. E. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109–1112 (1997).
- Grandbois, M., Beyer, M., Rief, M., Clausen-Schaumann, H. & Gaub, H. E. How strong is a covalent bond? *Science* **283**, 1727–1730 (1999).
- Mehta, A. D., Rief, M., Spudis, J. A., Smith, D. A. & Simmons, R. M. Single-molecule biomechanics with optical methods. *Science* **283**, 1689–1695 (1999).
- Kellermayer, M. S., Smith, S. B., Granzier, H. L. & Bustamante, C. Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science* **276**, 1112–1116 (1997).
- Keller, T. C. S. Molecular bungees. *Nature* **387**, 233–235 (1997).
- Tskhovrebova, L., Trinick, J., Sleep, J. A. & Simmons, R. M. Elasticity and unfolding of single molecules of the giant muscle protein titin. *Nature* **387**, 308–312 (1997).
- Al-Hassani, S. T. S. & Kaddour, A. S. Strain rate effects on GRP, KRP and CFRP composite laminates. *Key Eng. Mater.* **141**, 427–452 (1998).
- Greenwood, J. H. & Rose, P. G. Compressive behaviour of Kevlar 49 fibres and composites. *J. Mater. Sci.* **9**, 1809–1814 (1974).
- Lu, H., Israelowitz, B., Krammer, A., Vogel, V. & Schulten, K. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.* **75**, 662–671 (1998).
- Slater, G. W., Hubert, S. J. & Nixon, G. I. Construction of approximate entropic forces for finitely extensible nonlinear elastic (FENE) polymers. *Macromol. Theory Simulat.* **3**, 695–704 (1994).

Acknowledgements. We thank F. Lange and E. Kramer for help in understanding polymers, and H. Gaub and his group members both for helping to teach us how to unfold proteins, and for suggesting that organic components of abalone shells might be a good system in which to study polymer unfolding. This work was supported by the NSF (P.K.H.), by NSF through the Materials Research Laboratory, the Office of Naval Research, the Army Research Office MURI program, and the NOAA National Sea Grant College Program.

Correspondence and requests for materials should be addressed to B.L.S. (e-mail: bettye@physics.ucsb.edu).

Chiral nematic order in liquid crystals imposed by an engineered inorganic nanostructure

K. Robbie*†, D. J. Broer‡ & M. J. Brett*

* Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, T6G 2G7, Canada

‡ Philips Research Laboratories, Prof. Holstlaan 4, 5656, AA Eindhoven, The Netherlands

Control over the orientational order of liquid crystals (LCs) is critical to optical switching and display applications. Porous polymer networks have been used to influence the orientation of embedded chiral liquid crystals¹, yielding for example reflective displays. Here we show that inorganic films with a porous structure engineered on the submicrometre scale by glancing-angle deposition^{2,3} can be used to control the orientation of LCs impregnated into the voids. The inorganic material contains helical columns that orient rod-like nematic LCs into a phase similar to a chiral nematic^{1,4} but with direct control of the local molecular arrangement (for example, the helical pitch) imposed by the inorganic microstructure. We also show that reactive LC molecules in this composite material can be crosslinked by photopolymerization while retaining the imposed structure.

Thin films deposited at an oblique angle onto rotating substrates were first investigated in 1959⁵, and were found to display optical activity at one optical wavelength. In 1992, Azzam⁶ described a number of optical properties and devices that might be expected to result from thin films with a helical variation in optical anisotropy, based on oblique deposition (see, for example, refs 7, 8) onto rotating substrates. Robbie and Brett⁹ found in 1994 that porous films could be fabricated with submicrometre control of the constituent columnar structure. With motivation from theoretical investigations by Lakhtakia and Weiglhofer¹⁰, we then fabricated thin films with porous helical structures and demonstrated optical activity over a range of optical wavelengths^{11–13}. The films exhibit optical rotation and circular dichroism as would be predicted from comparisons with other chiral systems, including isotropic chiral mediums such as solutions of chiral molecules¹⁴, and chiral LCs¹⁵. This technique, which we named glancing-angle deposition (GLAD), was then used to produce a wide range of porous thin-film structures with various materials^{2,16}, was expanded to allow greater structural control³, and was investigated for use with sputtered films¹⁷.

Liquid crystals embedded in porous networks have been studied extensively in recent years because of their important present role, and promising future, in electro-optic switching and display technologies¹. Confining LCs in small-scale porous structures significantly affects the molecular ordering, and response to external fields, of the LC molecules, allowing electro-optic properties to be tailored. In most of this work, LCs are confined to organic polymeric networks formed either by phase separation or by emulsification. Inorganic porous networks, such as porous aerogels, Anopore membranes (aluminium oxide membranes containing cylindrical honeycomb pores) and porous glasses, have also been investigated. Twisted orientational structures have been observed in LCs confined in submicrometre spherical¹⁸ and cylindrical¹⁹ cavities, but elastic energy prevents multiple rotations of the molecular helix without adding chiral additives.

We deposited porous films with helical columns of MgF₂ onto glass substrates (Corning 7059) using GLAD^{2,12}. A detailed description of the deposition process is given in ref. 2. The films (Fig. 1a) are composed of helical columns of MgF₂ with 15 turns and a helical pitch of ~350 nm (the pitch of a helix is the distance along the helical axis for one period of revolution of the helix). The density of similar films has been previously measured with a microbalance mass measurement technique², and found to be ~60% of the bulk density. The accuracy of this measurement is limited, as only similar samples were measured.

The porous chiral film was impregnated with various materials including: water, an optically isotropic polymer (SR349, Sartomer, West Chester, Pennsylvania), and reactive (C3M) and non-reactive (ZLI4792 and E7) nematic LCs (Merck). To impregnate the film, the sample was covered with a thin glass slide held in place with small pieces of tape at two edges. A small drop of the impregnating material was placed at one of the open edges and rapidly filled the porous film, as observed by a change in light scattering.

To investigate the optical response of chiral films impregnated with LCs, circular polarized spectroscopic optical transmission was measured. The light switching property of chiral (or cholesteric) liquid crystals (CLCs), circular Bragg reflection, occurs between wavelengths $\lambda_1 = pn_o$ and $\lambda_2 = pn_e$ where n_o and n_e are the ordinary

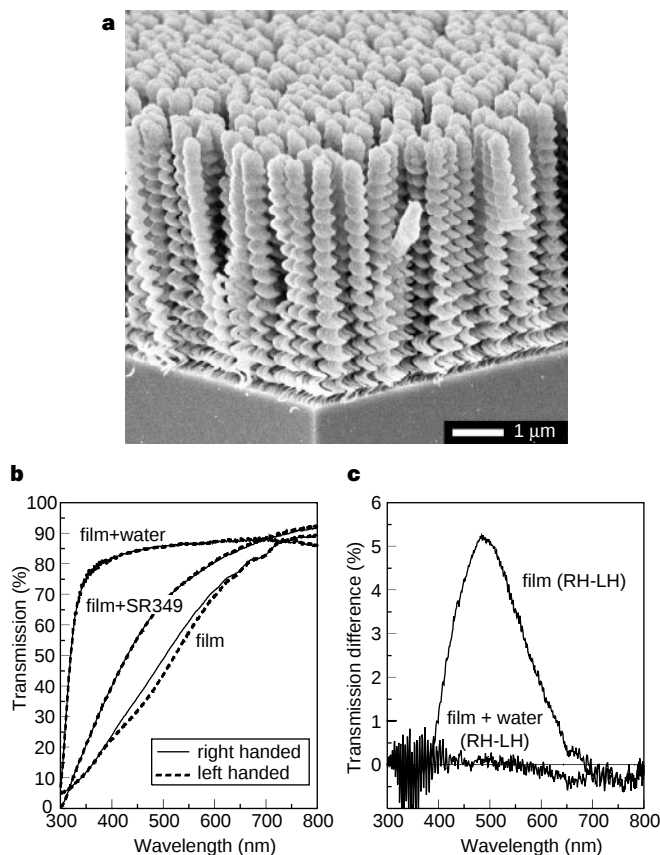


Figure 1 Porous inorganic GLAD film impregnated with optically isotropic fluids. **a**, Scanning electron micrograph of as-deposited GLAD thin film of MgF₂ ($n = 1.38$) on glass with 15 helical turns of pitch ~350 nm. **b**, Absolute transmission through the film and substrate for right-handed (RH) and left-handed (LH) circularly polarized light for the as-deposited film, the film impregnated with water ($n = 1.33$), and the film impregnated with an optically isotropic polymer-SR349 ($n = 1.56$ in bulk), cured with ultraviolet exposure. The difference between the 'film + water' and 'film' transmission spectra is attributed to diffuse scattering from inhomogeneities within the film. Weak transmission below 350 nm is due to absorption in the glass substrate. **c**, Difference spectrum of transmitted right-handed minus transmitted left-handed light for the as-deposited film and with a water impregnate.

† Present address: Department of Physics, Queen's University, Kingston, Ontario, K7L 3N6, Canada.

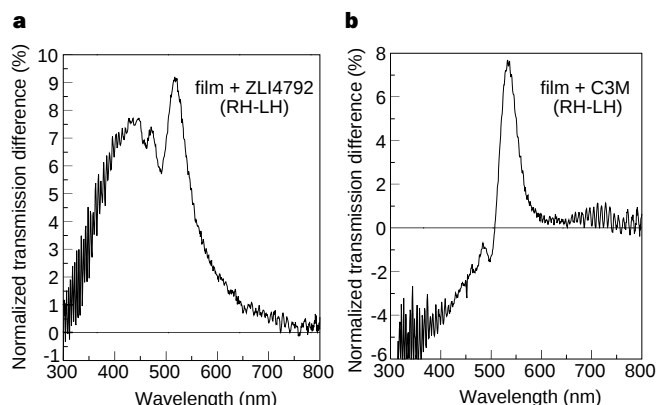


Figure 2 Transmission of circularly polarized light through a porous inorganic GLAD film with LC impregnates ZLI4792 and C3M. Difference spectra are shown for right-handed minus left-handed transmitted light through the film shown in Fig. 1a impregnated with **a**, a non-reactive LC blend, ZLI4792 ($n_o = 1.479$ and $n_e = 1.573$), and **b**, a LC diacrylate, C3M ($n_o = 1.5488$ and $n_e = 1.6880$) polymerized at 80 °C. The difference has been normalized to the transmittance of the right-handed circularly polarized light to minimize diffuse-scattering contributions. Absolute transmission was comparable to that observed for the film impregnated with SR349 (Fig. 1b), with the ZLI4792 impregnate exhibiting slightly higher transmission and the C3M exhibiting slightly lower transmission.

and extraordinary refractive indices of the locally uniaxial structure, and p is the pitch of the helical structure. Within this reflection band, right-handed circularly polarized light is reflected from a right-handed helix, and light with left-handed circular polarization is transmitted. Alternatively, left-handed circularly polarized light is reflected from a left-handed helix, and light with right-handed circular polarization is transmitted. Wavelengths outside the reflection band are transmitted in all polarizations. The width of the reflection band is less than ~ 100 nm for typical materials with $n_o \approx 1.5$ and $n_e \approx 1.7$. An approximate expression for the reflectivity R can be derived⁴:

$$R(\lambda) = \frac{k^4 \delta^2 \sin^2(\beta L)}{4q_0^2 \beta^2 + k^4 \delta^2 \sin^2(\beta L)} \quad (1)$$

where

$$\begin{aligned} \beta^2 &= k^2 + q_0^2 - k\sqrt{4q_0^2 + k^2\delta^2}, \\ k &= \frac{\pi(n_o + n_e)}{\lambda}, \quad \delta = \frac{n_e^2 - n_o^2}{n_e^2 + n_o^2}, \quad q_0 = \frac{2\pi}{p} \end{aligned} \quad (2)$$

The peak theoretical reflectivity can be very high for thicker layers of CLCs; up to $R \approx 99.988\%$ (a contrast ratio of ~ 40 dB) for a 5- μm -thick layer of a typical CLC material with $n_o \approx 1.5$ and $n_e \approx 1.7$. For thinner, or less than ideally aligned, layers, the observed reflectivity can be considerably less. High contrasts have been difficult to realize in bulk CLCs because of loss of alignment with thicker layers. In most display applications, however, very high contrast is not necessary, and the need for low switching voltages makes thinner LC layers more desirable. In other optical switching applications, such as for optical communications, higher contrast than that obtainable with bulk or polymer stabilized LCs is required. A more accurate numerical method for describing Bragg reflection in CLC layers is given by Berreman²⁰ and was used recently^{21,22}.

The circularly polarized transmission through the film shown in Fig. 1a without any impregnate is shown in Fig. 1b and c. Figure 1b shows the transmission through the film/substrate system ('film'), as measured. Figure 1c is the difference spectrum where the left-handed transmission is subtracted from the right-handed transmission to show the differential effect of the film on the two circular polarizations. The transmission difference has a peak at ~ 480 nm

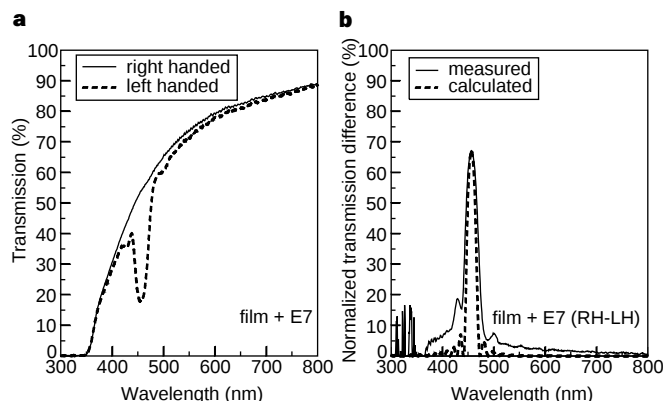


Figure 3 Transmission of circularly polarized light through a porous inorganic GLAD film with LC impregnate E7, and comparison with theory. Absolute **(a)** and difference **(b)** transmission spectra are shown for a GLAD film similar to that shown in Fig. 1a impregnated with E7 LC ($n_o = 1.5211$ and $n_e = 1.7476$). In **b**, a calculated spectrum is shown assuming chiral nematic alignment of the LC within the voids of the GLAD film.

which corresponds (within experimental uncertainty) with the estimated peak wavelength of $\lambda = pn_{\text{avg}} = 350 \times 1.2 = 420$ nm, where n_{avg} is an effective refractive index of the film which is estimated as the density (60%) weighted sum of the refractive index of the MgF_2 columns ($n = 1.38$) and the index of the air in the voids ($n = 1.00$). The slightly low prediction of peak wavelength could be a result of some water absorption within the voids which would tend to increase the effective index and shift the peak wavelength to higher values.

When the film was impregnated with water ($n = 1.33$, chosen to closely match the index of MgF_2 , $n = 1.38$), all transmission difference was lost, as is shown in Fig. 1b and c ('film + water'). The difference spectrum (Fig. 1c) shows some structure with differences less than 1%, but this may be due to imperfections in the spectrometer configuration. The second optically isotropic impregnating material tested was an ethoxylated bisphenol diacrylate (SR349), an optically isotropic polymer that has an index of 1.56 after curing with ultraviolet exposure. The transmission spectra are shown in Fig. 1b ('film + SR349'). As with water, no significant difference between right- and left-handed transmission is seen ($< 1\%$). Also, transmission reduction from diffuse scatter is less than found for the film alone, but not as low as for water. This is attributed to the larger index mismatch between the polymer and the film MgF_2 . From the transmission spectra through the film with and without the optically isotropic impregnates, we conclude that the chiral optical response seen in the as-deposited film arises from scattering at the film/air interfaces of the helix-shaped columns of MgF_2 . When the voids are filled with an optically isotropic index-matching material, the effects disappear.

When optically anisotropic impregnates such as LCs are used, there is a significant effect. An area of the film shown in Fig. 1a was filled with the reactive LC diacrylate C3M and polymerized at 80 °C, and another area was filled with the non-reactive LC blend ZLI4792. Another MgF_2 film, with a slightly larger pitch, was impregnated with a non-reactive nematic LC, E7. All three of these LC materials usually exhibit nematic phases at room temperature, and as such have no differential effect on circularly polarized light. The difference spectra for the films impregnated with ZLI4792 and C3M are shown in Fig. 2. If the LC impregnates are induced into a chiral nematic order by the helical structure of the GLAD films, the circular Bragg reflection will have a maximum at $\lambda = pn_{\text{avg}}$. Taking a simple mixing rule of the density-weighted sum of the index of MgF_2 and the average index of the LCs $\lambda_{\text{C3M}} = pn_{\text{avg}} = 350(0.6 \times 1.38 + 0.4(1.5488 + 1.6880)/2) = 520$ nm, the maxima should occur at ~ 520 nm for C3M and 500 nm for

ZLI4792. Measured transmission difference maxima agree within experimental error, confirming that chiral GLAD films induce a type of chiral order in simple nematic LCs, similar to the chiral nematic phase seen in CLCs, and with the pitch matching the physical pitch of the GLAD helical columns.

Absolute transmission and normalized transmission difference through the film impregnated with the nematic E7 are shown in Fig. 3. Here the observed transmission difference is much stronger than that observed with the C3M and ZLI4792 impregnates. Again, the maxima occurs (within experimental error) at the wavelength predicted assuming chiral nematic order of the LC within the GLAD film. Figure 3b shows a calculated curve from equation (1) where the transmission difference peak magnitude and wavelength have been chosen to match the measured spectrum. In matching the measured curve, an average index of $n_{\text{avg}} = 1.43$ and an anisotropy $\delta = 0.024$ were used. The simple mixing rule described above predicts an average index of $n_{\text{avg}} = 1.47$, again within experimental error. The optical anisotropy is much less than that of E7 alone ($\delta \approx 0.14$), but as the hybrid is composed of 60% isotropic MgF_2 , and 40% LC, this is not unexpected. It is quite difficult to estimate the anisotropy for the composite material and this was not attempted.

Differential transmission of circularly polarized light through the hybrid thin-film/LC material shows that some chiral ordering of the LC is induced by the helical backbone of the porous film fabricated with GLAD. However, a true single-domain 'planar' chiral nematic texture, where all LC molecules lie parallel to the substrate with the helical axis perpendicular, is not the only possible molecular orientational texture. The observed differences in effect with varying material (E7 produced much stronger Bragg reflection than the other LCs), as well as the low anisotropy needed to match theory to observations in Fig. 3b, suggest that a more complicated alignment in fact occurs. One peculiarity of the LC E7 is that there is a cyano group at the end of the rod-like molecules. Dipole interactions of this group with polar surfaces (such as glass) make these molecules susceptible to alignment perpendicular to the surface (homeotropic alignment). It is therefore likely that at the surfaces of the MgF_2 helices, the E7 molecules will align homeotropically. But in the larger voids between the helical MgF_2 columns, the complex mix of alignments arising from homeotropic alignment onto helical columns must resolve to reduce elastic energies. One possible texture is a tilted orientation that follows the rotation of the nearby helices. The 'bulk' void regions would have a chiral texture that imitates the nearby helices, but with a tilt to reduce elastic energy and eliminate domain boundaries. Another possibility would be a tube-like structure with homeotropic alignment (perpendicular to the substrate) within the 'bulk' void regions. This texture is more like that seen in the absence of a helical backbone, where all molecules are aligned homeotropically (perpendicular to the substrate). The MgF_2 columns would then introduce small but periodic variations in the local orientation of the LC near the columns. These two possible textures probably describe the limits of a continuum of alignments; the alignment that occurred would depend on the inorganic film material and surface treatments, LC chemistry, temperature and applied fields. Both possibilities would help to explain the low observed anisotropy of the E7 mixture, as the full anisotropy of the tilted LC is not seen for light propagating along the helical axis.

As an example, and as a comparison to conventional LC techniques, one of us (D.J.B.) has shown that the small bandwidth of transmission/reflection (usually <100 nm) of a CLC layer can be increased by producing a structure with a graded pitch. This was accomplished by crosslinking reactive CLCs by photopolymerization of a diffusion-controlled concentration gradient of CLC²³. This produced a solidified structure, however, and does not allow for switching. The GLAD technique allows LC alignment structures, such as graded pitches, to be engineered for a desired optical response, and will still allow switching of the LC. GLAD films

composed of multiple layers with different pitches, handedness and structure have been fabricated, and will be investigated in future work. □

Received 26 February; accepted 5 May 1999.

1. Crawford, G. P. & Zumer, S. (eds) *Liquid Crystals in Complex Geometries* (Taylor and Francis, London, 1996).
2. Robbie, K. & Brett, M. J. Sculptured thin films and glancing angle deposition: growth mechanics and applications. *J. Vac. Sci. Technol. A* **15**, 1460–1465 (1997).
3. Robbie, K., Sit, J. C. & Brett, M. J. Advanced techniques for glancing angle deposition. *J. Vac. Sci. Technol. B* **16**, 1115–1122 (1998).
4. Belyakov, V. A. & Dmitrienko, V. E. Optics of chiral liquid crystals. *Sov. Sci. Rev. A* **13**, (1989).
5. Young, N. O. & Kowal, J. Optically active fluorite films. *Nature* **183**, 104–105 (1959).
6. Azzam, R. M. A. Chiral thin solid films: method of deposition and applications *Appl. Phys. Lett.* **61**, 3118–3120 (1992).
7. Hodgkinson, I. J. & Wu, Q. H. *Birefringent Thin Films and Polarizing Elements* (World Scientific, Singapore, 1997).
8. van Kranenburg, H. & Lodder, J. C. Tailoring growth and local composition by oblique-incidence deposition: a review and new experimental data. *Mater. Sci. Eng. R* **11**, 295–354 (1994).
9. Robbie, K., Friedrich, L. J., Dew, S. K., Smy, T. & Brett, M. J. Fabrication of thin films with highly porous microstructures. *J. Vac. Sci. Technol. A* **13**, 1032–1035 (1995).
10. Lakhtakia, A. & Weiglhofer, W. S. On light propagation in helicoidal bianisotropic mediums. *Proc. R. Soc. Lond. A* **448**, 419–437 (1995).
11. Robbie, K., Brett, M. J. & Lakhtakia, A. First thin film realization of a helicoidal bianisotropic medium. *J. Vac. Sci. Technol. A* **13**, 2991–2993 (1995).
12. Robbie, K., Brett, M. J. & Lakhtakia, A. Chiral sculptured thin films. *Nature* **384**, 616 (1996).
13. Rovira, P. I. et al. Rotating-compensator multichannel transmission ellipsometry of a thin-film helicoidal bianisotropic medium. *Thin Solid Films* **313–314**, 373–378 (1998).
14. Barron, L. D. *Molecular Light Scattering and Optical Activity* (Cambridge Univ. Press, 1982).
15. Chandrasekhar, S. *Liquid Crystals* (Cambridge Univ. Press, 1977).
16. Robbie, K., Hnatiw, A. J. P., Brett, M. J., MacDonald, R. I. & McMullin, J. N. Inhomogeneous thin film optical filters fabricated using glancing angle deposition. *Electron. Lett.* **33**, 1213–1214 (1997).
17. Sit, J. C., Vick, D., Robbie, K. & Brett, M. J. Thin film microstructural control using glancing angle deposition by sputtering. *J. Mater. Res.* **14**, 1197–1199 (1999).
18. Xu, F., Kitzrow, H.-S. & Crooker, P. P. Electric-field effects on nematic droplets with negative dielectric anisotropy. *Phys. Rev. E* **46**, 6535–6540 (1992).
19. Ondris-Crawford, R. J. et al. Surface molecular anchoring in microconfined liquid crystals near the nematic-smectic-A transition. *Phys. Rev. E* **48**, 1998–2005 (1993).
20. Berreman, D. W. Optics in stratified and anisotropic media: 4×4 -matrix formulation. *J. Opt. Soc. Am.* **62**, 502–510 (1972).
21. St John, W. D., Fritz, W. J., Lu, Z. J. & Yang, D.-K. Bragg reflection from cholesteric liquid crystals. *Phys. Rev. E* **51**, 1191–1198 (1995).
22. Xu, M., Xu, F. & Yang, D.-K. Effects of cell structure on the reflection of cholesteric liquid crystal displays. *J. Appl. Phys.* **83**, 1938–1944 (1998).
23. Broer, D. J., Lub, J. & Mol, G. N. Wide-band reflective polarizers from cholesteric polymer networks with a pitch gradient. *Nature* **378**, 467–469 (1995).

Acknowledgements. This work was supported in part by the Natural Sciences and Engineering Council of Canada and the Alberta Microelectronic Centre.

Correspondence and requests for materials should be addressed to K.R. at Queen's University (e-mail: robbie@physics.queensu.ca).

A reversibly antigen-responsive hydrogel

Takashi Miyata, Noriko Asami & Tadashi Uragami

Chemical Branch, Faculty of Engineering and High Technology Research Center, Kansai University, Suita, Osaka, 564-8680, Japan

Stimuli-responsive hydrogels that undergo abrupt changes in volume in response to external stimuli such as pH, temperature and solvent composition have potential applications in biomedicine and the creation of 'intelligent' materials systems, for example as media for drug delivery, separation processes and protein immobilization. Hydrogels have been reported that respond to pH^{1–3}, temperature^{4–13}, electric fields^{14–16} and saccharides^{17–22}. For some biomedical applications it would be very useful to have a material whose swelling response was dictated by a specific protein. Here we report such a material, which swells reversibly in a buffer solution in response to a specific antigen. The hydrogel was prepared by grafting the antigen and corresponding antibody to the polymer network, so that binding between the two introduces crosslinks in the network. Competitive binding of the free antigen triggers a change in gel volume owing to breaking of these non-covalent crosslinks. In addition, we show that the hydrogel displays shape-memory behaviour, and that stepwise changes in antigen concentration can induce pulsatile permeation of a

protein through the network.

Our strategy is to use the reversible binding between an antigen and an antibody as the crosslinking mechanism in a semi-interpenetrating network (semi-IPN) hydrogel (Fig. 1a). The hydrogel can swell in the presence of a free antigen because the intra-chain antigen-antibody binding can be dissociated by exchange of the grafted antigen for free antigen. In the absence of a free antigen, the

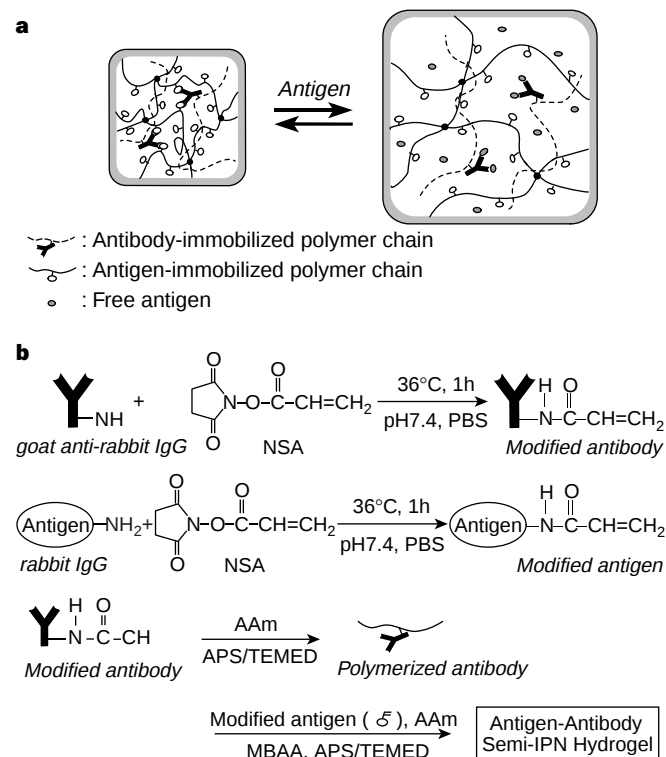


Figure 1 Strategy for the preparation of an antigen-responsive hydrogel. **a**, Diagram of a suggested mechanism for the swelling of an antigen-antibody semi-IPN hydrogel in response to a free antigen. **b**, Synthesis of the antigen-antibody semi-IPN hydrogel. The GAR IgG was chemically modified by coupling it with *N*-succinimidylacrylate (NSA) in phosphate buffer solution, using the method reported by Hoffman *et al.*²³. NSA (0.4 mg) was added to a phosphate buffer solution (0.02 M, pH 7.4) containing the GAR IgG (100 mg) (NSA/GAR IgG molar ratio is 6:1), and the reaction was then incubated at 36 °C for one hour to introduce the vinyl groups into the GAR IgG. The resultant vinyl(GAR IgG) was purified by gel filtration, and a phosphate buffer solution containing the vinyl(GAR IgG) was obtained at a concentration of 1.91 mg ml⁻¹. After acrylamide (AAm) (30 mg) was added to 570 mg of the vinyl(GAR IgG) solution, together with 0.01 ml of 0.1 M aqueous ammonium persulphate (APS) and 0.01 ml of 0.8 M aqueous *N,N,N',N'*-tetramethylethylenediamine (TEMED) as redox initiators, the copolymerization was performed at 25 °C for 3 h to synthesize the polymerized GAR IgG. As the relative molecular mass of the resultant polymerized GAR IgG was ~220,000, the average number of IgG incorporated into a polymerized GAR IgG was ~1. Furthermore, vinyl(rabbit IgG) was also synthesized by modifying rabbit IgG with NSA in the same manner as the vinyl(GAR IgG). The resultant vinyl(rabbit IgG) (2.46 mg), AAm (82 mg) and *N,N'*-methylenebisacrylamide (MBAA) (0.1 wt% relative to AAm) as a crosslinker were dissolved in 600 mg of phosphate buffer solution containing the polymerized GAR IgG to form the antigen-antibody binding. As soon as 0.01 ml of 0.1 M aqueous APS and 0.01 ml of 0.8 M aqueous TEMED were added into the mixture as redox initiators, it was injected into a glass capillary tube with an inner diameter of 3 mm (or into a mould membrane preparation), and the polymerization was performed at 25 °C for 3 h. After the polymerization, the resultant antigen-antibody semi-IPN hydrogels were immersed in phosphate buffer to remove any residual chemicals and unreacted monomers. Polyacrylamide (PAAm) semi-IPN hydrogels were also prepared as a reference material by the redox copolymerization of AAm and MBAA in the presence of linear PAAm in the same manner.

hydrogel shrinks. We used rabbit immunoglobulin G (rabbit IgG) and goat anti-rabbit IgG (GAR IgG) as the antigen and antibody, respectively. The GAR IgG was purified with an affinity column in which rabbit IgG was immobilized; the binding of those IgGs formed the stimuli-responsive crosslinking point in the hydrogel. The typical method to prepare antigen-antibody semi-IPN hydrogel is shown in Fig. 1b. The rabbit IgG and GAR IgG were chemically modified by coupling them with *N*-succinimidylacrylate (NSA in Fig. 1b) in phosphate buffer solution, using the method reported by Hoffman *et al.*²³. The copolymerization of the resultant vinyl(GAR IgG) with acrylamide (AAm) was performed with redox initiators to synthesize the polymerized GAR IgG. An antigen-antibody semi-IPN hydrogel was prepared by the copolymerization of the vinyl(rabbit IgG), acrylamide and *N,N'*-methylenebisacrylamide (MBAA) as a cross-linker in the presence of the polymerized GAR IgG in a glass capillary tube.

The antigen-responsive swelling of the antigen-antibody and polyacrylamide semi-IPN hydrogels was examined after their swelling attained equilibrium in 0.2 M phosphate buffer solution (pH 7.4); the water content of the equilibrated hydrogels was 97 wt%. The swelling ratio of the antigen-antibody semi-IPN hydrogel increased abruptly following the addition of free rabbit IgG. Furthermore, the swelling ratio of the polyacrylamide semi-IPN hydrogel was independent of the antigen concentration in the buffer solution, but the swelling ratio of the antigen-antibody semi-IPN hydrogel was strongly dependent upon it (Fig. 2). This indicates that the latter hydrogel is rabbit-IgG-responsive, and has an antigen-sensing function.

To determine the mechanism responsible for the rabbit-IgG-responsive swelling, complex formation between the antigen and antibody in the hydrogel was investigated using the BIAcore system^{24,25} (Pharmacia Biotec, Inc.). GAR IgG was immobilized on a sensor tip surface in the BIAcore system; the surface plasmon resonance induced at the interface between the surface and the solution containing modified rabbit IgG was then examined to determine the kinetic parameter for the complex formation. As a result, the binding constants of the GAR IgG to the native and polymerized rabbit IgGs were $8.90 \times 10^8 \text{ M}^{-1}$ and $2.61 \times 10^8 \text{ M}^{-1}$, respectively. This suggests that the binding of the GAR IgG with rabbit IgG polymerized in the antigen-antibody hydrogel was much weaker than to native rabbit IgG, due to the denaturation of the modified rabbit IgG. Therefore, adding free native rabbit IgG can induce the dissociation of the complex between GAR IgG and rabbit IgG grafted to the semi-IPN hydrogel.

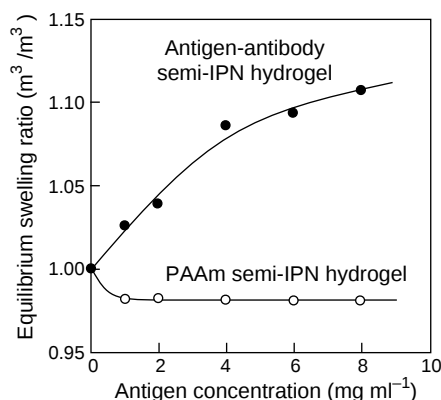


Figure 2 Effects of the free antigen concentration on the hydrogel swelling ratio. Shown is the equilibrium swelling ratio of the PAAm semi-IPN hydrogel (open circles) and the antigen-antibody semi-IPN hydrogel (filled circles) in phosphate buffer solution at 25 °C as a function of antigen concentration in the buffer solution. The swelling ratio of the hydrogels was determined by the ratio of their changing diameters, which were measured in the solution using an optical microscope.

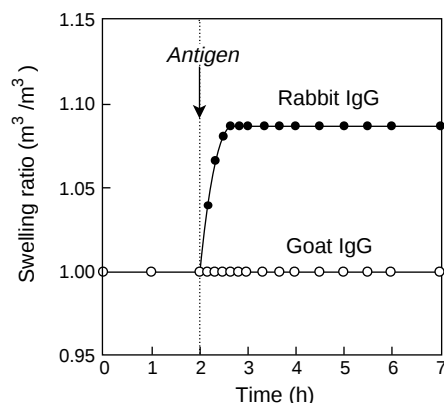


Figure 3 Antigen recognition by the antigen-antibody semi-IPN hydrogel. The swelling ratio of the antigen-antibody semi-IPN hydrogel is shown, before and after the addition of goat IgG (open circles) and rabbit IgG (filled circles). Measurements were made after the swelling had attained equilibrium in phosphate buffer solution at 25 °C. The concentration of the antigen in the phosphate buffer solution was 4 mg ml⁻¹.

The crosslinking density of a hydrogel can be determined by measuring its compressive modulus^{10,18,26}. We performed compressive modulus measurements to examine the changes in the crosslinking density of the hydrogel following the addition of free antigen (see Supplementary Information). The crosslinking density decreased gradually in proportion to the increasing free rabbit IgG concentration in the phosphate buffer solution, indicating that the antigen-responsive swelling changes are due to decreasing crosslinking density, which is consistent with the mechanism depicted in Fig. 1a. We note that the denaturation of the rabbit IgG in preparing the antigen-responsive hydrogel was an important factor in the exchange between native and polymerized rabbit IgGs.

The swelling ratio of the hydrogel did not change following the addition of goat IgG, whereas addition of rabbit IgG led to a drastic increase (Fig. 3). These results mean that the hydrogel can recognize a specific antigen and change its structure chemomechanically. As the GAR IgG in the hydrogel forms a complex with rabbit IgG only due to the excellent antigen-recognition of this antibody, the addition of goat IgG results in no dissociation of the complex between the rabbit IgG and the GAR IgG grafted to the hydrogel.

Most potential applications of the antigen-antibody semi-IPN hydrogel require that the swelling be reversible. We investigated the reversibility in response to stepwise changes in the antigen concentration, as well as the consequent permeation of a model protein drug (haemoglobin) through a membrane fabricated from the hydrogel. Figure 4 shows the swelling ratio of the polyacrylamide and antigen-antibody semi-IPN hydrogels, and permeation profiles of a model protein drug through their membranes, as a function of time; during this time, they were alternately immersed in phosphate buffer with and without 4 mg ml⁻¹ of rabbit IgG. There was no change in the swelling ratio of the polyacrylamide semi-IPN hydrogel following its alternate immersion in phosphate buffer with and without rabbit IgG. However, the antigen-antibody semi-IPN hydrogel swelled immediately in the presence of rabbit IgG, and shrank gradually in its absence. In contrast, another antigen-antibody hydrogel without a semi-IPN structure also swelled following the addition of free rabbit IgG, but did not shrink following immersion in phosphate buffer without rabbit IgG (data not shown). This suggests that the semi-IPN structure is important in the reversibility of the antigen-responsive swelling changes. In the semi-IPN hydrogel the grafted GAR IgG is trapped in a network containing grafted rabbit IgG, and so the hydrogel can shrink reversibly because the crosslinks between grafted GAR IgG and grafted rabbit IgG reform in phosphate buffer (Fig. 1a). Furthermore, because the hydrogel was prepared by copolymerizing with

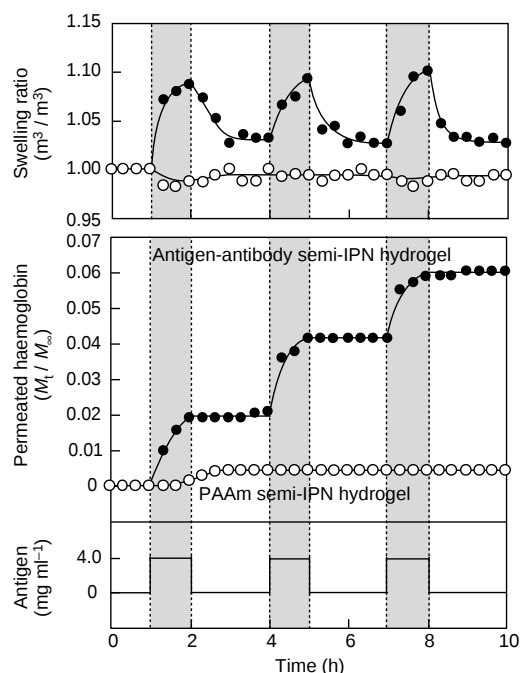


Figure 4 Reversible swelling changes and antigen-responsive permeation profiles. Shown are swelling ratio changes of the PAAm semi-IPN hydrogel (open circles) and the antigen-antibody semi-IPN hydrogel (filled circles), and a permeation profile of a model drug through their membranes, in response to stepwise changes in the antigen concentration between 0 and 4 mg ml⁻¹ at 25 °C. Permeation experiments were performed at 25 °C with magnetic stirring, using a glass cell of two parts separated by a hydrogel membrane. One chamber of the cell was filled with phosphate buffer solution containing haemoglobin (1 mg ml⁻¹) as a model protein drug, and other chamber was filled with a phosphate buffer solution in which the antigen concentration was changed stepwise between 0 and 4 mg ml⁻¹. M_t indicates the amount of the model drug permeated through the hydrogel membrane at a given time; M_∞ indicates the total amount of the model drug in the chamber of the cell.

vinyl(rabbit IgG), which forms a complex with the polymerized GAR IgG, the resultant semi-IPN hydrogel can 'memorize' the shape of its initial semi-IPN structure. Thus, hydrogel shrinks in the absence of rabbit IgG on the basis of a 'shape-memory-like' effect. In addition, it is apparent from Fig. 4 that the model drug permeates through the membrane in the presence of rabbit IgG but is stopped in its absence, suggesting that this approach might permit drug delivery in response to a specific antigen. □

Received 19 October 1998; accepted 19 April 1999.

1. Tanaka, T. Collapse of gels and the critical endpoint. *Phys. Rev. Lett.* **40**, 820–823 (1978).
2. Tanaka, T. *et al.* Phase transitions in ionic gels. *Phys. Rev. Lett.* **45**, 1636–1639 (1980).
3. Siegel, R. A. & Firestone, B. A. pH-dependent equilibrium swelling properties of hydrophobic polyelectrolyte copolymer gels. *Macromolecules* **21**, 3254–3259 (1988).
4. Hirokawa, Y. & Tanaka, T. Volume phase transition in a nonionic gel. *J. Chem. Phys.* **81**, 6379–6380 (1984).
5. Hoffman, A. S. Applications of thermally reversible polymers and hydrogels in therapeutics and diagnostics. *J. Controlled Release* **6**, 297–305 (1987).
6. Bae, Y. H., Okano, T., Hsu, R. & Kim, S. W. Thermo-sensitive polymers as on-off switches for drug release. *Makromol. Chem. Rapid Commun.* **8**, 481–485 (1987).
7. Bae, Y. H., Okano, T. & Kim, S. W. A new thermo-sensitive hydrogel: interpenetrating polymer networks from N-acryloylpyrrolidine and poly(oxyethylene). *Makromol. Chem. Rapid Commun.* **9**, 185–189 (1988).
8. Chen, G. & Hoffman, A. S. Graft copolymers that exhibit temperature-induced phase transitions over a wide range of pH. *Nature* **373**, 49–52 (1995).
9. Yoshida, R. *et al.* Comb-type grafted hydrogels with rapid de-swelling response to temperature changes. *Nature* **374**, 240–242 (1995).
10. Miyata, T., Nakamae, K., Hoffman, A. S. & Kankazaki, Y. Stimuli-sensitivities of hydrogels containing phosphate groups. *Macromol. Chem. Phys.* **195**, 1111–1120 (1994).
11. Aoki, T. *et al.* Temperature-responsive interpenetrating polymer networks constructed with poly(acrylic acid) and poly(N,N-dimethylacrylamide). *Macromolecules* **27**, 947–952 (1994).
12. Stayton, P. S. *et al.* Control of protein-ligand recognition using a stimuli-responsive polymer. *Nature* **378**, 472–474 (1995).
13. Feil, H., Bae, Y. H., Feijen, J. & Kim, S. W. Molecular separation by thermosensitive hydrogel membranes. *J. Membrane Sci.* **64**, 283–294 (1991).

14. Tanaka, T., Nishio, I., Sun, S.-T. & Ueno-Nishio, S. Collapse of gels in an electric field. *Science* **218**, 467–469 (1982).
15. Kwon, I. C., Bae, Y. H. & Kim, S. W. Electrically erodible polymer gel for controlled release of drugs. *Nature* **354**, 291–293 (1991).
16. Osada, Y., Okuzaki, H. & Hori, H. A polymer gel with electrically driven motility. *Nature* **355**, 242–244 (1992).
17. Kokufuta, E., Zhang, Y.-Q. & Tanaka, T. Saccharide-sensitive phase transition of a lectin-loaded gel. *Nature* **351**, 302–304 (1991).
18. Miyata, T., Jikihara, A., Nakamae, K. & Hoffman, A. S. Preparation of poly(glucosyloxyethyl methacrylate)-concanavalin A complex hydrogel and its glucose-sensitivity. *Macromol. Chem. Phys.* **197**, 1135–1146 (1996).
19. Lee, S. J. & Park, K. Synthesis and characterization of sol-gel phase-reversible hydrogels sensitive to glucose. *J. Mol. Recogn.* **9**, 549–557 (1996).
20. Obidat, A. A. & Park, K. Characterization of glucose dependent gel-sol phase transition of polymeric glucose-concanavalin A hydrogel system. *Pharm. Res.* **13**, 989–995 (1996).
21. Obidat, A. A. & Park, K. Characterization of protein release through glucose-sensitive hydrogel membranes. *Biomaterials* **18**, 801–806 (1997).
22. Kataoka, K. *et al.* Totally synthetic polymer gels responding to external glucose concentration: their preparation and application to on-off regulation of insulin release. *J. Am. Chem. Soc.* **120**, 12694–12695 (1998).
23. Shoemaker, S. G., Hoffman, A. S. & Priest, J. H. Synthesis and properties of vinyl monomer/enzyme conjugates. *Appl. Biochem. Biotech.* **15**, 11–24 (1987).
24. Natsume, T. *et al.* Interactions between collagen-binding stress protein HSP47 and collagen. *J. Biol. Chem.* **269**, 31224–31228 (1994).
25. Murai, N., Taguchi, H. & Yoshida, M. Kinetic analysis of interactions between GroEL and reduced α -lactalbumin. *J. Biol. Chem.* **270**, 19957–19963 (1995).
26. Flory, P. J. *Principles of Polymer Chemistry* (Cornell Univ. Press, New York, 1953).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Correspondence and requests for materials should be addressed to T.M. (e-mail: tmiyata@ipcku.kansai-u.ac.jp).

The use of path integration to guide route learning in ants

B. Schatz*, S. Chameron†, G. Beugnon† & T. S. Collett*

* Sussex Centre of Neuroscience, School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK

† Laboratoire d'Éthologie et de Psychologie Animale, CNRS-UMR 5550, Université Paul-Sabatier, Toulouse cedex 04, France

Cataglyphid ants travelling between their nest and feeding site follow familiar routes along which they are guided by views of the surrounding landscape^{1–3}. On bare terrain, with no landmarks available, ants can still navigate using path integration⁶. They continually monitor their net distance and direction from the nest, so that they can return home from any point using their computed 'home vector'⁷. Here we ask whether path integration also provides signals to reinforce the learning of visual landmarks. A fall in the value of the home vector indicates when a homing ant moves in roughly the correct direction, and that it is appropriate to store those views that can guide subsequent trips to the nest⁸. We tested this hypothesis by training the ant *Cataglyphis cursor* to negotiate a variety of mazes that led from a feeding site back to the nest. Efficient passage of each maze required an ant to discriminate between different pairs of shapes⁹. We show that if the value of the home vector drops while the ant approaches and passes a shape, the shape's appearance is learnt, but if the vector grows, or is absent, no visual learning occurs. Path integration may both help ants navigate through an unfamiliar landscape, and assist them to become familiar with it.

Cataglyphis cursor is a mediterranean ant with a mean foraging range of 6 m (ref. 10) and it readily forages in the laboratory. Our first maze (Fig. 1a) was designed so that the home vector decreased as the ant travelled towards the nest from the feeder. Foragers from the nest reached a sucrose feeder through an unobstructed tunnel (122 cm). Individually marked foragers were taken from the feeder to a waiting box and then transferred singly to the start of the maze. The maze was directly above the tunnel and aligned with it. It had four compartments connected by passages. Each compartment had one entrance and two exits, of which one was blocked and the other led to the next compartment (Fig. 1a). Open and closed exits were

labelled by black shapes on a white background (Fig. 1a, inset). To prevent ants from negotiating the maze by following a fixed path, we changed the side of the open exit and its associated shape after each trial. Ants learn which visual shape signals the open exit in each compartment and their sequence along the maze⁹.

Four other combinations of maze and tunnel were devised in which the home vector increased, or was absent, over all or part of the ant's homeward passage through the maze. Instead of a linear tunnel and maze, the tunnel and/or the maze were bent into a U shape, so introducing a large discrepancy between the ant's direction and that of its home vector (Fig. 1b–e). Our aim was to test whether visual learning is impaired under these conditions.

The combined acquisition curves of a group of ants trained to reach food through the linear tunnel and to home along the linear maze are plotted separately for each compartment (Fig. 2a). Each plot cumulates, over successive trials, the mean number of errors made by each ant on its first choice of exit on each trial. With no learning, the curves should stick to the line indicating chance performance. In fact, the curves deviate significantly ($P < 0.01$) downwards from that line after 13, 11, 5 and 1 trials in compartments 1 to 4, respectively. During training, there was attrition in the number of ants, as shown by the numbers given at the top of each column. Of the 16 ants remaining at the end of the 100-trial observation period, 13 had learnt compartments 1 and 2, 16 had learnt compartment 3, and 15 had learnt compartment 4. We show in the Methods that the 1-trial score in compartment 4 can be explained by the ants' sensitivity to the alternating pattern used in training. Thus, the gradient in error scores over the four compartments may, to some extent, be a consequence of the ants' intrinsic motor strategies.

Visual learning was absent when the directions of the home vector and the maze differed. We trained ants to reach the feeder through a U-shaped tunnel, so that the presumed home vector was very short and, if directed at all, was perpendicular to the maze. The return to the nest was either through a U-shaped (Fig. 1b) or through a linear

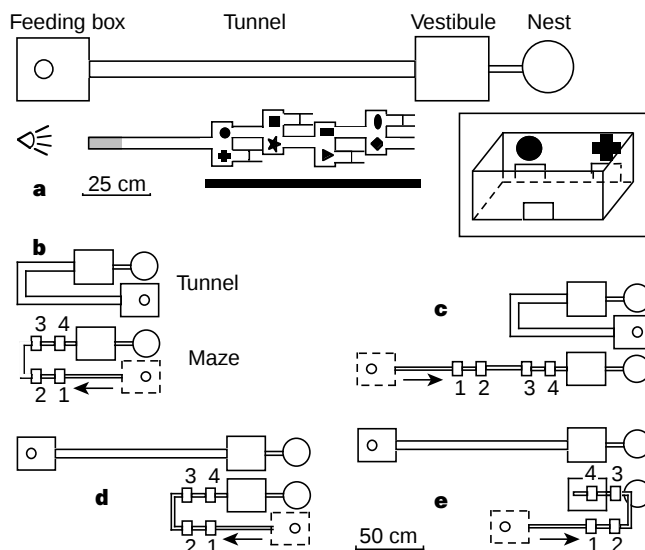


Figure 1 Diagrams of mazes used to test visual learning. In all conditions, the maze is placed directly over the tunnel. **a**, Ants approach the feeder along a linear tunnel and return home singly through a linear maze. The open exit from compartment 1 of the maze is labelled by a black circle, and the open exits from the other three compartments by a star, a rectangle and a diamond, respectively. Directional cues come from a floodlight at one end of the room and a large piece of black card fixed to one wall. Inset, three-dimensional sketch of one compartment to show ant's view of the shapes and exits. **b**, Tunnel and maze with same compartments and visual shapes as in **a** are both bent into a U shape. **c**, U-shaped tunnel and linear maze. **d**, **e**, Linear tunnel followed by a U-shaped maze. Upper calibration refers to **a**, lower calibration refers to **b–e**.

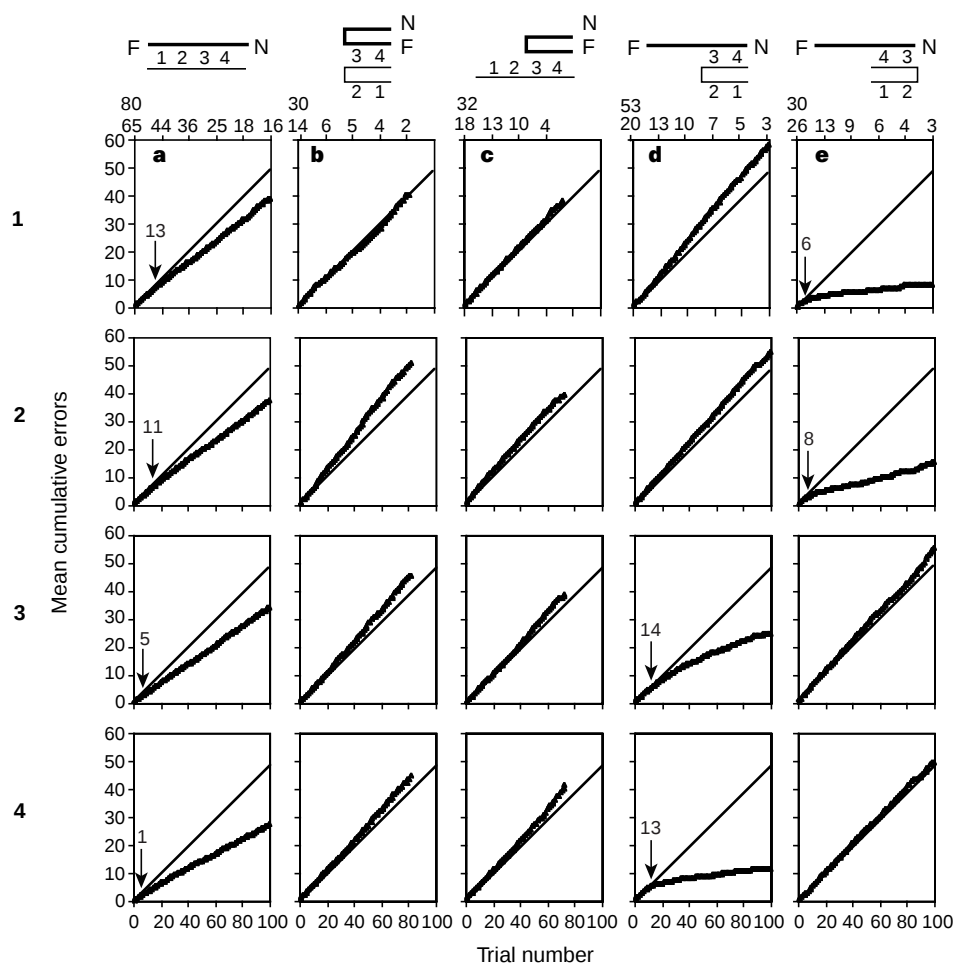


Figure 2 Acquisition curves for the five arrangements of tunnels and mazes shown in Fig. 1. Each column (a–e) shows cumulative error plots for one maze-tunnel arrangement, as displayed at the top of the column, where the thickened line represents the tunnel and the thin line the maze, with the numbers representing the four compartments. N and F signify the nest and feeder respectively. The data in the first column are from ants in which training was interrupted periodically to test for visual-sequence learning⁹. The head of each

column shows the number of marked ants; the upper number gives the total at the start of the experiment, the other numbers give the sample size every 20 trials. Each row gives the error plots for one compartment (1–4). Mean number of errors per ant are accumulated over successive trials. Only the first choice made in each box on each trial is considered, so that there is a 50% chance of making an error in each compartment. The diagonal line indicates random choice. Arrows show when the plots depart from random choice at $P < 0.01$.

maze (Fig. 1c). Ants were initially reluctant to travel through the maze, and many ants had to be eliminated from the experiment because of their excessive slowness, using criteria which we had established earlier⁹. Cumulative error plots did not depart significantly downwards from the line showing chance performance, and no individual learnt to distinguish the shapes (Fig. 2b, c). In the next series of experiments, ants reached the feeder through a linear tunnel and returned home through a U-shaped maze (Fig. 1d, e). The home vector is aligned with the maze in one arm, but points in the opposite direction to the maze along the other arm. Under these conditions, the same ants learnt the shapes in compartments in which the directions of the home vector and the maze agreed, but showed no learning if these directions were opposed (Fig. 2d, e).

Ants returning from an outward trip through the linear tunnel update their home vector, as demonstrated by their ability to compensate for an unexpected detour¹¹. Single ants were taken from the feeder to a tunnel, 106 cm long, that was orientated at 45° either to the left or to the right of the outward tunnel. Once an ant reached the end of the 45° tunnel, it was placed in the centre of a circular arena that was kept in the same room position for both tests. Ants walked towards the periphery in a direction that compensated for the enforced detour. The mean direction with respect to the homeward tunnel was 93° (95% confidence

interval $\pm 12^\circ$, $n = 32$) to the right for leftward detours, and it was 88° (95% confidence interval $\pm 13^\circ$, $n = 35$) to the left for rightward detours, close to the required 90°. The results presented in Fig. 2b are puzzling, if it is supposed that ants also update their home vector on returning to the nest after an outward journey through the U-shaped tunnel. In this case, one might expect ants to learn the shapes in compartments 3 and 4. We suggest that the feeder at the end of the U-shaped tunnel is so close to home that ants do not compute a home vector. The ants in the maze are searching for rather than aiming for home, and we suppose that visual learning is not engaged during their search.

At the end of training with the configuration of Fig. 1d, when errors in compartments 3 and 4 were infrequent, we transferred the shapes that usually labelled compartment 4 to compartment 1. Ants in this test failed to pick the entrance marked by the positive shape (50%, $n = 24$) indicating that they ignore even familiar visual cues if the directions of their path and home vector are discrepant. Whereas, with no discrepancy, positive shapes are recognized out of their usual context⁹.

Despite the absence of visual learning in ants reaching the food through the U-shaped tunnel, their overall performance improved. Mean travel time through the maze became faster as training continued, with a corresponding drop in the total number of

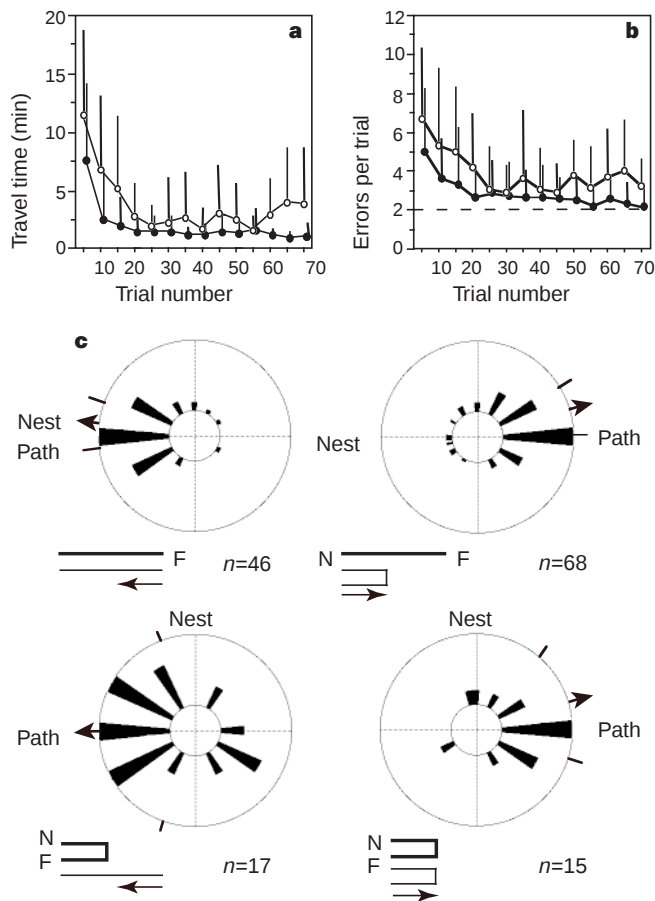


Figure 3 'Motor' learning without a compatible home vector. **a, b**, Indicators of improved performance in ants trained with U-shaped tunnels (Fig. 1c, d). **a**, Travel time per individual from its entry to compartment 1 to its exit from the maze is plotted over successive trials. Times are averaged over blocks of five trials. Numbers on abscissa indicate trial number of the last trial in the block. Open circles (U-shaped tunnel and U-shaped maze) and filled circles (U-shaped tunnel and linear maze) show mean duration and bars show the standard deviation of the mean. **b**, Total number of errors per individual on successive trials. All of the ants' entries to wrong exits, including revisiting errors, are counted and averaged over blocks of five trials. Dashed horizontal line shows theoretical performance assuming no revisiting errors and no visual learning. **c**, Radial plots of ants in circular arena (diameter 20cm) after training in different tunnel/maze arrangements as sketched below each circular histogram. Thick and thin lines denote outward and homeward paths respectively with arrow indicating initial homeward direction. The labelling around the histogram indicates the direction of the nest from the feeding box (nest) and the initial direction of the homeward path (path). Arrows show mean direction flanked by 95% confidence intervals, *n* indicates the number of ants tested in each condition.

erroneous choices per trial (Fig. 3a, b). The error measure here includes revisiting errors, in contrast to that used for Fig. 2. The optimal performance of a shape-blind ant should asymptote at 2.0 errors per trip. Increased speed and fewer errors indicate that ants have acquired the habit of travelling from one compartment to the next, even if they must walk without their home vector, and that they have developed a strategy to limit revisiting errors. It seems unlikely, therefore, that the lack of visual learning can be attributed solely to the ants' distress on being forced to move without the aid of a home vector. We conclude that the results shown in Fig. 2 are evidence for a link between path integration and visual acquisition. However, the details of the link are not yet clear. Thus, an alternative interpretation of our results might be that learning occurs when the ant's body is aligned with the home vector, rather than when the home vector diminishes.

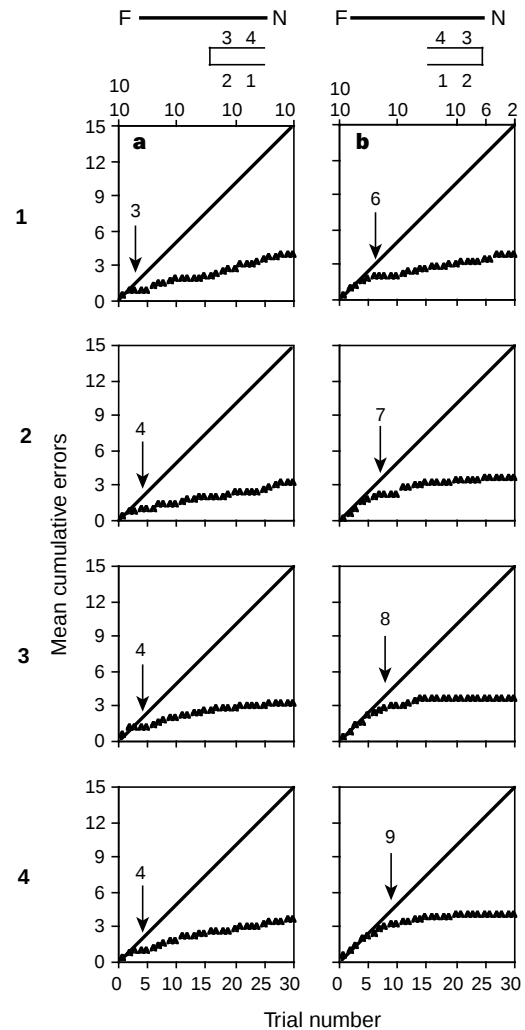


Figure 4 Acquisition curves for fixed-path learning. Ants reached the food along a linear tunnel and returned through one or other version of the U-shaped maze. The path through the maze alternated between left and right in successive compartments. Errors scored and conventions as in Fig. 2.

The acquisition of 'motor' skills seems to follow different rules, and it can occur in situations in which visual learning fails. This is illustrated by the following two experimental tests. First, ants learned to walk in the direction in which they were funnelled by the corridor at the start of their return trip, even if it was against their home vector. Ants trained to approach the feeder along a linear or U-shaped tunnel and to return along either a linear or U-shaped path were taken from the feeder and placed in the centre of a circular arena suspended above the feeder. The histograms in Fig. 3c show the 30° sector in which ants were located on reaching the wall of the arena. Mean vectors all point roughly in the initial direction of the ant's accustomed homeward path, irrespective of the direction of the nest. Second, ants learnt fixed paths without the aid of a compatible home vector. Ants arriving at the feeder, through a linear tunnel, were trained in a U-shaped maze. In one maze, the first arm pointed against the home vector and the second arm along it. The open exits and the shapes in each compartment were kept on the same side on every trial, so that ants could ignore the shapes and learn a fixed path through the maze, alternating sides as they proceeded from compartment to compartment. After very few trials, the ants' performance was almost errorless in all four compartments (Fig. 4a). Acquisition of a fixed path was nearly as rapid in the second U-shaped maze (Fig. 4b) where the first arm

pointed along the home vector and the second arm in the opposite direction. A final test (not illustrated) showed that ants were equally quick at learning a fixed path if they had to choose the same side in all compartments.

The ability to learn fixed paths rapidly, without a compatible home vector, indicates that ants may temporarily store a sequence of path segments on every trip home. If path segments can be encoded economically as local vectors⁵, the ant's speculative learning of a path should not make unreasonable demands on its short-term memory. In contrast, ants seem to avoid the speculative learning of views. Instead, they use home vectors to trigger or permit the storage of significant scenes. A reinforcement signal emanating from the home vector would mean that different scenes viewed along a route could all be acquired on the same trip independently of each other. When learning a complex route, an ant may depend initially upon the rapid acquisition of path segments through motor learning. Routes may then be stabilized through visual learning that is guided by path integration. By this means paths can become as economical as the terrain allows, with acquired motor patterns moulded to the improved route. □

Methods

Maze training. Experiments were done using queenright colonies of *Cataglyphis cursor* maintained in the laboratory⁹. At the start of an experiment, some foragers were taken from the sucrose feeder at the end of the tunnel and marked with paint. Marked individuals always returned home through the maze. They were collected on their arrival at the feeder and released in a waiting box with another similar feeder. Ants taken from the waiting box were placed singly at the start of the maze. Each ant performed between five and ten trips a day. During visual training, positive and negative shapes were exchanged after each trial, and the compartments repositioned so that the previously blocked exit was open and the previously open exit blocked. Training thus involved an alternating motor pattern from trial to trial (L(left)R(right)LR or RLRL through the exits in the four compartments). To examine how this pattern influenced the ant's choice, we performed two tests in which ants that had foraged through the linear tunnel, but were naive to the maze, performed a single homeward passage through the linear maze. In test 1, the open exits from compartments 1 to 4 were LLLR or RRRL, and in test 2 the open compartments were LRLR or RLRL. In both tests 1 and 2, ants chose randomly in compartment 1. In compartment 2, they tended to choose the exit on the side that had been open in compartment 1. The behaviour of the two groups diverged in compartments 3 and 4. Ants in test 1 persisted in choosing the exit on the side that had been open in compartments 1 and 2. Thus, in compartment 4, 22 ants chose the negative shape (and the closed exit) and 3 ants chose the positive shape (binomial test, two-tail $P < 0.001$). In contrast, the ants in test 2 tended to follow the alternating pattern. In compartment 4, 35 ants chose the positive shape and 12 ants chose the negative shape (binomial test, two-tail $P < 0.001$). This sensitivity to the alternating training sequence is likely to reduce the errors made in the last two compartments during visual acquisition. Although cataglyphid ants do not lay chemical trails¹², it is important with fixed-path learning to ensure that ants cannot be guided by chemical cues. Compartments and tunnels were therefore switched to distribute any potential chemical cues evenly over the maze.

Tests in the arena. Ants were carried to the circular arena (diameter 20 cm) on a straw which was placed upright in a hole in the centre. The ant climbed off the straw and walked towards the periphery. We recorded in which 5° or 30° sector the ant was located when it reached the wall of the arena.

Statistical treatment. On each trial we scored the choice of exits made in each compartment and the time spent travelling from the first compartment to the end of the maze. The binomial distribution was applied to the group plots of Figs 2 and 4, using a one-tailed test, to show when the error score dropped below chance at $P < 0.01$. To assess the learning of individual ants contributing to the group plots in Fig. 2, we counted the number of errors made by each ant in its first choice of exit in each compartment. We then used the binomial test to determine whether an individual's score within a compartment was less than that attributable to chance ($P < 0.05$). Mean directions and confidence intervals in Fig. 3c were obtained as described by Batschelet¹³.

Received 29 December 1998; accepted 23 April 1999.

1. Wehner, R., Harkness, R. D. & Schmid-Hempel, P. Foraging strategies in individually searching ants, *Cataglyphis bicolor* (Hymenoptera: Formicidae). *Akad. Wiss. Lit. Mainz, Math.-Naturwiss. Kl.* **1**, 1–79 (1983).
2. Wehner, R. Spatial organization of foraging behaviour in individually searching desert ants, *Cataglyphis* (Sahara desert) and *Ocymyrmex* (Namib desert). *Experientia* **54** (Suppl.), 15–42 (1987).
3. Collett, T. S., Dillmann, E., Giger, A. & Wehner, R. Visual landmarks and route following in desert ants. *J. Comp. Physiol. A* **170**, 435–442 (1992).
4. Wehner, R., Michel, B. & Antonsen, P. Visual navigation in insects: coupling of egocentric and geocentric information. *J. Exp. Biol.* **199**, 129–140 (1996).
5. Collett, M., Collett, T. S., Bisch, S. & Wehner, R. Local and global vectors in desert ant navigation. *Nature* **394**, 269–272 (1998).
6. Mittelstaedt, H. & Mittelstaedt, M. L. In *Avian Navigation* (eds Papi, F. & Wallraff, H. G.) 290–297 (Springer, New York, 1982).
7. Wehner, R. & Wehner, S. Insect navigation: use of maps or Ariadne's thread? *Ethol. Ecol. Evol.* **2**, 27–48 (1990).
8. Collett, T. S. Rapid navigational learning in insects with a short lifespan. *Connection Sci.* **10**, 255–270 (1998).
9. Chameron, S., Schatz, B., Pastergue-Ruiz, I., Beugnon, G. & Collett, T. S. The learning of a sequence of visual patterns by the ant *Cataglyphis cursor*. *Proc. R. Soc. Lond. B* **265**, 2309–2313 (1998).
10. Lenoir, A., Nowbahari, E., Querard, L., Pondicq, N. & Delalande, C. Habitat exploitation and intercolonial relationships in the ant *Cataglyphis cursor* (Hymenoptera: Formicidae). *Acta Oecologica* **11**, 3–18 (1990).
11. Schmidt, I., Collett, T. S., Dillier, F.-X. & Wehner, R. How desert ants cope with enforced detours on their way home. *J. Comp. Physiol. A* **171**, 285–288 (1992).
12. Harkness, R. D. Quantitative observations on the foraging of nests of an ant (*Cataglyphis bicolor* F.) in Greece. *Acta Entomol. Jugosl.* **13**, 21–33 (1977).
13. Batschelet, E. *Circular Statistics in Biology* (Academic, London, 1981).

Acknowledgements. We thank V. Fourcassie, P. Graham, M. Jones and J. Lauga for discussion. Financial support came from the European Commission (TMR Program) to B.S. and T.S.C., the French Ministry of Research to S.C., the Science de la Cognition Program to G.B., and the UK BBSRC and the Human Frontier Science Program to T.S.C.

Correspondence and requests for materials should be addressed to T.S.C. (e-mail: T.S.Collett@sussex.ac.uk).

Hox genes in brachiopods and priapulids and protostome evolution

Renaud de Rosa^{†‡}, Jennifer K. Grenier^{†‡}, Tatiana Andreeva[§], Charles E. Cook^{||}, André Adoutte[†], Michael Akam^{||}, Sean B. Carroll[‡] & Guillaume Balavoine^{||¶}

[†]Laboratoire de Biologie Cellulaire 4, CNRS UPRESA Q8080, Université Paris-Sud, 91405 Orsay Cedex, France

[‡]Howard Hughes Medical Institute and Laboratory of Molecular Biology, University of Wisconsin, 1525 Linden Drive, Madison, Wisconsin 53706, USA

[§]Department of Embryology, Biological Institute, St Petersburg University, 199034 St Petersburg, Russia

^{||}Laboratory for Development and Evolution, University Museum of Zoology, Department of Zoology, Downing St, Cambridge CB2 3EJ, UK

[¶]These authors contributed equally to this work.

Understanding the early evolution of animal body plans requires knowledge both of metazoan phylogeny and of the genetic and developmental changes involved in the emergence of particular forms. Recent 18S ribosomal RNA phylogenies suggest a three-branched tree for the Bilateria comprising the deuterostomes and two great protostome clades, the lophotrochozoans¹ and ecdysozoans². Here, we show that the complement of Hox genes in critical protostome phyla reflects these phylogenetic relationships and reveals the early evolution of developmental regulatory potential in bilaterians. We have identified Hox genes that are shared by subsets of protostome phyla. These include a diverged pair of posterior (*Abdominal-B*-like) genes in both a brachiopod and a polychaete annelid, which supports the lophotrochozoan assemblage, and a distinct posterior Hox gene shared by a priapulid, a nematode and the arthropods, which supports the

[¶] Present address: Centre de Génétique Moléculaire, CNRS UPR 9061, Ave. de la Terrasse, Bâtiment 26, 91198 Gif-sur-Yvette, France.

ectoderm. The ancestors of each of these two major protostome lineages had a minimum of eight to ten Hox genes. The major period of Hox gene expansion and diversification thus occurred before the radiation of each of the three great bilaterian clades.

Hox genes have key roles in patterning the antero-posterior axis of bilaterians³. The reconstruction of the evolutionary history of the Hox gene family is thus central to understanding the evolution of bilaterian body plans and the relationship between genetic and morphological complexity. Hox gene distribution also provides a

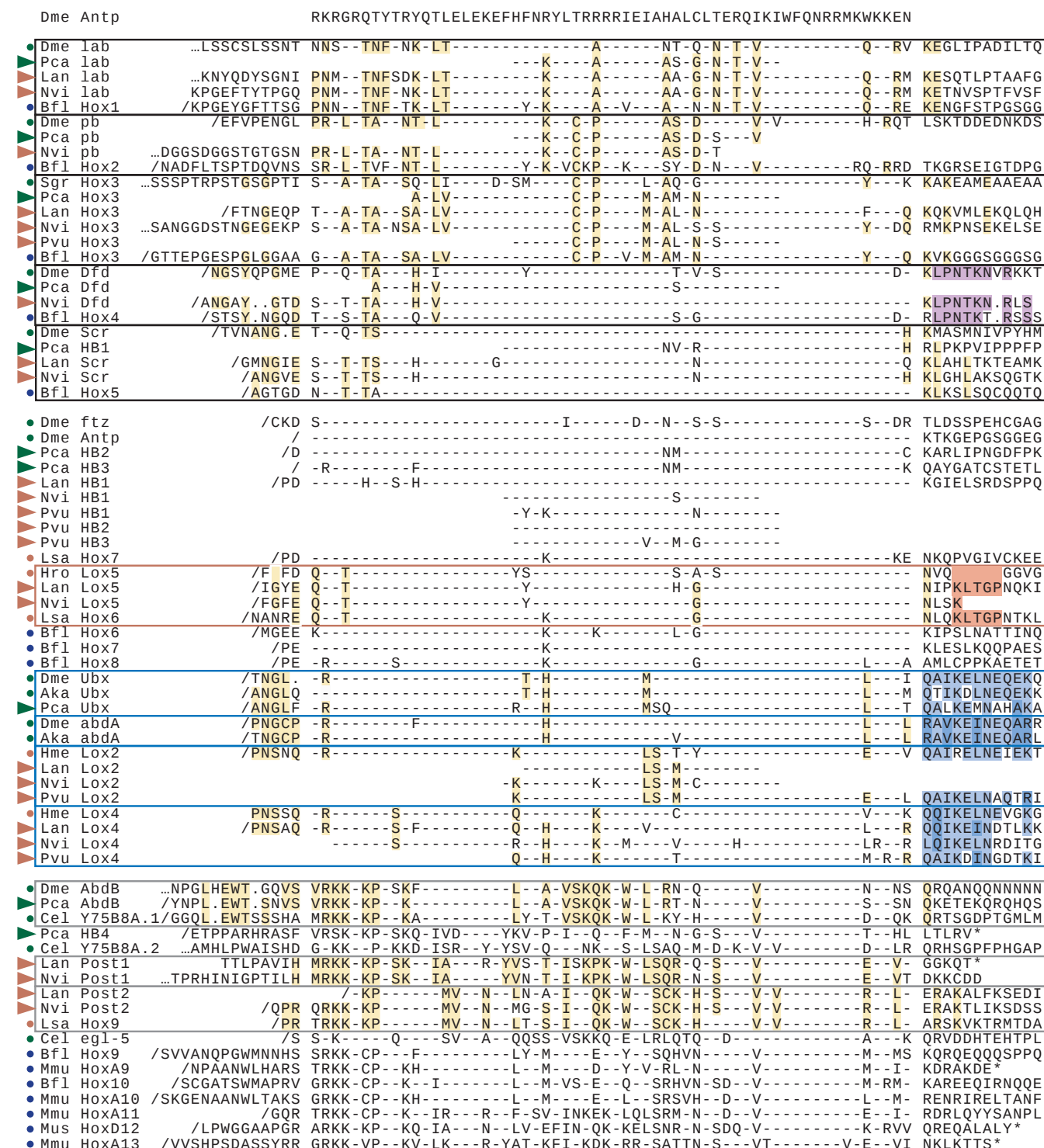


Figure 1 Alignment of Hox homeodomains and flanking sequences. Sequences first reported here are indicated by an arrowhead, aligned in comparison to previously published Hox genes. Colour code for arrowheads and dots: green, ecdysozoans; orange, lophotrochozoans; blue, deuterostomes. Within the homeodomain, dashes indicate identity to *Drosophila melanogaster Antp* (at top). Anterior and central genes that can be grouped by orthology across metazoan phyla are enclosed within black boxes (*Dfd* peptide motif highlighted in

pink). Sequences boxed in orange contain the 'Lox5' peptide motif⁴ (highlighted in orange) and sequences boxed in blue contain the 'Ubd-A' peptide motif⁴ (highlighted in blue). Posterior genes are boxed in grey when orthology is well supported by phylogenetic analysis (see Fig. 2). Yellow highlighting shows conserved residues (>50% majority) that differ from *Antp* within each group of orthologues. See Table 1 for species abbreviations.

Table 1 Species names and abbreviations used in Figures

Common name	Species name	Abbreviation	References
Amphioxus	<i>Branchiostoma floridae</i>	Bfl	26, 27
Ascidian	<i>Ciona intestinalis</i>	Cin	*
Brachiopod	<i>Lingula anatina</i>	Lan	This paper
Crustacean	<i>Artemia franciscana</i>	Afr	*
Flatworms	<i>Girardia tigrina</i>	Gti	*
	<i>Polycelis nigra</i>	Pni	*
Fruit fly	<i>Drosophila melanogaster</i>	Dme	18
Gastropod	<i>Patella vulgata</i>	Pvu	This paper
Leeches	<i>Helobdella robusta</i>	Hro	*
	<i>Helobdella triserialis</i>	Htr	*
	<i>Hirudo medicinalis</i>	Hme	*
Locust	<i>Schistocerca gregaria</i>	Sgr	*
Mouse	<i>Mus musculus</i>	Mmu	28
Nematode	<i>Caenorhabditis elegans</i>	Cel	19, 29
Nemertean	<i>Lineus sanguineus</i>	Lsa	14
Onychophoran	<i>Acanthokara kaputensis</i>	Aka	5
Polychaete	<i>Nereis virens</i>	Nvi	This paper
Priapulid	<i>Priapulus caudatus</i>	Pca	This paper
Sea urchins	<i>Strongylocentrotus purpuratus</i>	Spu	30
	<i>Heliocidaris erythrogramma</i>	Her	*
	<i>Triploneustes gratilla</i>	Tgr	*

* See Supplementary Information for complete reference list and accession numbers.

new tool for addressing long-standing issues in metazoan phylogeny^{4,5}. For both these purposes data need to be obtained from a broad sample of phyla, particularly those that help resolve the early branching pattern of the bilaterians. For example, if the proposed assemblages of lophotrochozoan phyla (lophophorates, annelids, molluscs, flatworms and others) and ecdysozoan phyla (arthropods, onychophorans, nematodes, priapulids and others) are correct^{1,2,6}, Hox gene identities ought to reflect this phylogeny. We have identified Hox genes from two previously unsampled phyla, a lophophorate (the brachiopod *Lingula anatina*) and a priapulid (*Priapulus caudatus*). Brachiopods were once thought to be deuterostomes⁷, whereas priapulids were often associated with the other ‘pseudocoelomates’ basal to the great majority of bilaterian phyla⁸. Both these phyla are among the most ancient bilaterians to appear in the Cambrian fossil record⁹. We have also collected an extensive Hox gene data set from a polychaete annelid (*Nereis virens*), and partial data from a gastropod mollusc (the limpet *Patella vulgata*). Many protostome Hox genes can clearly be assigned to one of five orthology groups that are also recognizable in deuterostomes: *labial (lab)/Hox1*, *proboscipedia (pb)/Hox2*, *zerknüllt (zen)/Hox3*, *Deformed (Dfd)/Hox4* and *Sex combs reduced (Scr)/Hox5*. These orthology groups are defined by conserved residues unique to each group that are shared across bilaterian phyla, and they are supported by phylogenetic analysis (Fig. 1 and Supplementary Information). The remaining protostome Hox genes cannot be identified as orthologues of specific deuterostome genes. Many can, however, be placed into groups on the basis of conserved amino acids within the homeodomain and distinct peptide motifs in the flanking sequences (Fig. 1) that are found in subsets of protostome phyla. In this way, we identify three conserved central (*Antennapedia*-like) genes shared by the brachiopod, the polychaete annelid and the gastropod mollusc that have been previously characterized in annelids and named *Lox5* (ref. 10), *Lox2* (ref. 12) and *Lox4* (refs 12, 13) (we continue to use these names until unambiguous linkage data allows rationalization of the nomenclature). Similarly, on the basis of conserved residues within and flanking the homeodomain, we group one of the priapulid central Hox genes with the *Ultrabithorax (Ubx)* gene. This gene has been described previously only in arthropods and onychophorans⁵. Additional central genes found in each organism cannot be clearly assigned to an orthology group because sequence divergence obscures relationships across phyla. Some of these may be new genes that have arisen from more recent gene duplication events. Previous studies of protostome Hox genes have identified at most

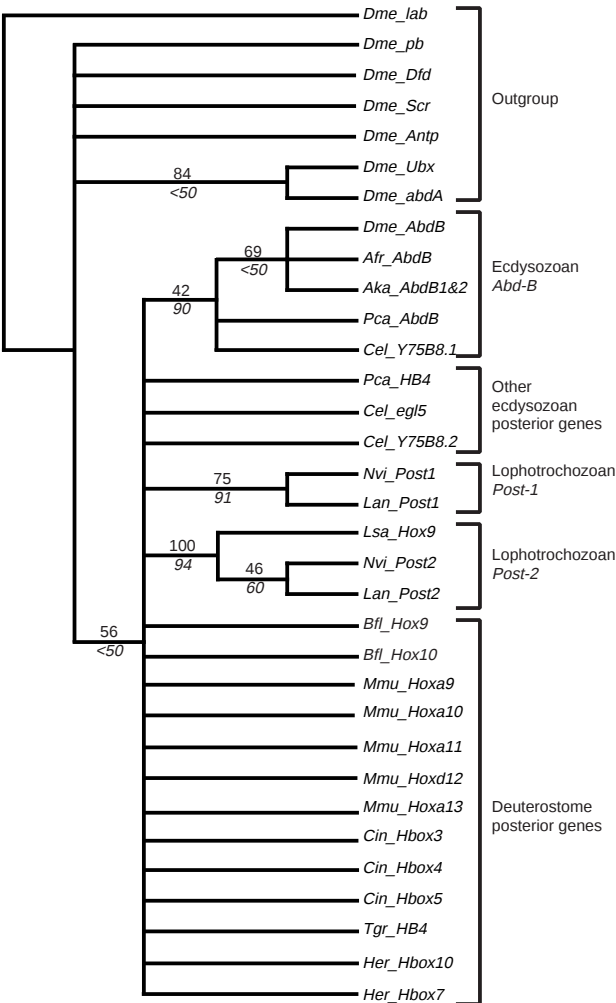


Figure 2 Phylogenetic analysis of bilaterian posterior Hox homeodomain sequences. Numbers above branches indicate per cent support in MP bootstrap analyses. Italicized numbers below branches indicate per cent reliability in a Puzzle ML tree. Note high support for the ecdysozoan *Abd-B* genes, the lophotrochozoan *Post-1* genes, and the lophotrochozoan *Post-2* genes. Nodes with low support (<50%) in all analyses have been collapsed to polytomies. See Supplementary information for phylogenetic analyses that include *caudal/cdx* genes.

a single posterior (*Abdominal-B*-like) gene^{5,14–18}, but we have found multiple posterior genes in several protostome phyla. We discovered two posterior genes in the lophophorate, the polychaete and the priapulid, and the *C. elegans* genome sequencing project has identified two previously uncharacterized posterior genes (in addition to *egl-5*) in this nematode¹⁹ (Fig. 1). On the basis of the homeodomain sequence, the brachiopod and polychaete genes fall into two new orthology groups that we term *Post-1* and *Post-2*; *Post-2* has also been found in the nemertean *Lineus sanguineus* (*LsHox9*; ref. 14) and a fragment of it in an oligochaete¹⁵. One of the priapulid posterior genes and one of the newly sequenced nematode genes (*Y75B8A.1*) are orthologous to the arthropod *Abdominal-B* (*Abd-B*) gene (Fig. 1). These orthology assignments are well supported by phylogenetic analyses of the homeodomain sequences (Fig. 2). However, none of the posterior Hox genes can be individually related to any of the deuterostome genes, as shown by the basal polytomy in Fig. 2; nor are any of these genes orthologous to the ParaHox *caudal/cdx* genes²⁰ (see Supplementary Information). Sequence divergence has obscured the origins of these posterior genes, and it is not possible to determine whether they arose from

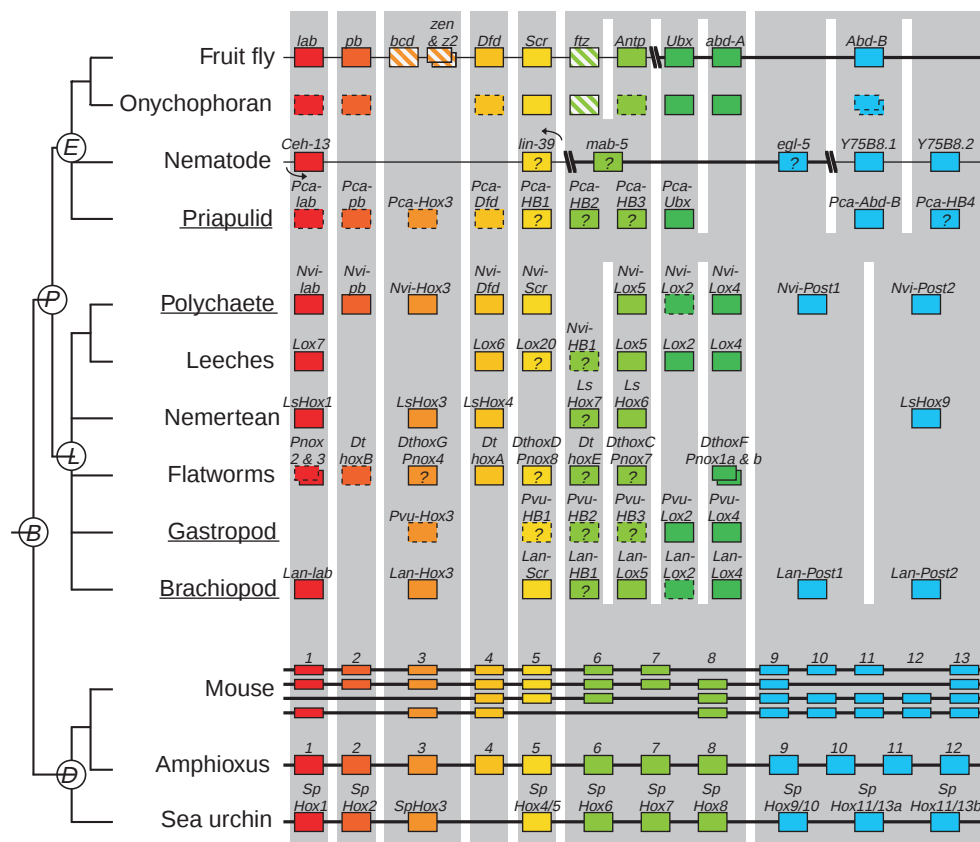


Figure 3 Distribution of Hox genes in bilaterians. The tree on the left summarizes the 18S rDNA phylogeny of Aguinaldo *et al.*². B, common bilaterian ancestor; D, common deuterostome ancestor; E, stem ecdysozoan; L, stem lophotrochozoan; P, common protostome ancestor. Horizontal black lines indicate mapping data (when available). Vertical white bars delineate orthologous genes or groups of

genes. Uncertain orthology relationships are indicated by question marks, dashed boxes indicate short PCR fragments, solid boxes indicate complete or almost complete homeoboxes, and striped colours indicate fast-evolving arthropod Hox sequences. The organisms studied in this paper are underlined.

independent duplications in each of the three clades, or whether multiple posterior Hox genes were inherited from a common bilaterian ancestor.

The identification of distinct Hox genes that are shared by some, but not all, protostome phyla constitutes independent molecular evidence for the division of protostomes into two great clades—the lophotrochozoans and the ecdysozoans. The animals within each of these lineages have central and posterior Hox genes specific to their clade, which provides a surprisingly strong signal of the deep division within the protostomes. The lophotrochozoans are characterized by the presence of *Lox5*, *Lox2*, *Lox4*, *Post-1* and *Post-2*, and the ecdysozoans by *Ubx* and *Abd-B*. Each of these genes can be identified by characteristic shared, derived residues in and near the homeodomain, which became fixed before the radiation of the crown phyla of each lineage. Not all of the proposed ecdysozoans definitively have all of the ‘ecdysozoan’ Hox genes (most notably *C. elegans*), but the unambiguous identification of an ‘ecdysozoan’ Hox gene in any animal, such as the *C. elegans* *Abd-B* gene (Y75B8A.1), supports the inclusion of that animal within the ecdysozoans. Thus, these shared sets of Hox genes provide signatures for each protostome clade, and can be used to infer the phylogenetic placement of other protostome phyla.

A comparison of the Hox genes that are shared by descendants of a particular lineage enables estimation of the number of Hox genes present in earlier animal ancestors. On the basis of shared Hox genes, we conclude that the stem lophotrochozoan (Fig. 3, Node L) had at least ten Hox genes and the stem ecdysozoan (Fig. 3, node E) at least eight. Following the same logic, we infer that the common protostome ancestor (Fig. 3, node P) also had at least eight Hox

genes and the common bilaterian ancestor (Fig. 3, node B) had at least seven (*lab/Hox1*, *pb/Hox2*, *Hox3*, *Dfd/Hox4*, *Scr/Hox5*, one additional central gene and one posterior gene). Indeed, seven Hox genes is a minimum estimate for the bilaterian ancestor, and the true number might have been higher. As more phyla have been sampled, we are struck by the fact that all clusters (except for the degenerate *C. elegans* cluster) contain nine or more genes. This indicates that most or all of the Hox genes that are present in extant bilaterians (Fig. 3) may have been present in the common ancestor, but that some orthology relationships have become obscured. If this new ‘early expansion’ hypothesis is true, then at least ten Hox genes could have existed in the protostome ancestor (*lab/Hox1*, *pb/Hox2*, *Hox3*, *Dfd/Hox4*, *Scr/Hox5*, three to four central genes and at least two posterior genes), or even more if the deuterostome condition of multiple posterior Hox genes is ancestral. The subsequent bilaterian history of Hox genes would have been primarily one of functional divergence and gene loss, rather than gene duplication. Regardless of the exact number of Hox genes in the bilaterian ancestor, the major period of progressive expansion of the Hox cluster due to tandem duplication events²¹ predated the radiation that generated the bilaterian crown phyla, concurrent with radical evolutionary changes in body architecture and development. □

Methods

L. anatina DNA was made available by the authors of a previous study²². *N. virens* specimens were collected in the Marine Biological Station of St Petersburg University, Chupa Inlet, White sea, Russia. *P. vulgata* specimens were collected in the Station Biologique, Roscoff, France. *P. caudatus* specimens were collected from Lopez Sound, Washington, USA and from Kasitsna Bay, Alaska,

USA. Short homeobox fragments were amplified by PCR from genomic DNA using degenerate primers (see Supplementary Information for details). Homeodomain sequences were extended using either ligation mediated-PCR^{5,23} or inverse-PCR¹⁷. Phylogenetic trees were built from aligned complete homeodomain sequences using unweighted maximum parsimony (MP) with PAUP*4.0b1 (ref. 24) and maximum likelihood (ML) with Puzzle²⁵. The MP analysis was performed with the following settings: heuristic search over 100 bootstrap replicates, MAXTREES set to 10,000 due to computer limitations, other parameters set to default values. The ML analysis was performed using the quartet puzzling tree search procedure with 10,000 puzzling steps and using a gamma-distributed model of rate heterogeneity with eight categories.

Received 11 February 1999; accepted 20 April 1999.

- Halanych, K. M. *et al.* Evidence from 18S ribosomal DNA that the lophophorates are protostome animals. *Science* **267**, 1641–1643 (1995).
- Aguinaldo, A. M. *et al.* Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489–493 (1997).
- Carroll, S. B. Homeotic genes and the evolution of arthropods and chordates. *Nature* **376**, 479–485 (1995).
- Balavoine, G. Are platyhelminthes coelomates without a coelom? An argument based on the evolution of *Hox* genes. *Am. Zool.* **38**, 843–858 (1998).
- Grenier, J. K., Garber, T. L., Warre, R., Whittington, P. M. & Carroll, S. Evolution of the entire arthropod *Hox* gene set predated the origin and radiation of the onychophoran/arthropod clade. *Curr. Biol.* **7**, 547–553 (1997).
- Adoutte, A., Balavoine, G., Lartillot, N. & de Rosa, R. Animal evolution: the end of the intermediate taxa? *Trends Genet.* **15**, 104–108 (1999).
- Nielsen, C. *Animal Evolution* (Oxford Univ. Press, Oxford, 1995).
- Lorenzen, S. in *The Origins and Relationships of Lower Invertebrates* (ed. Conway-Morris, S.) 210–223 (Clarendon, Oxford, 1985).
- Boardman, R. S., Cheetham, A. H. & Rowell, A. J. *Fossil Invertebrates* (Blackwell Scientific, Oxford, 1987).
- Shankland, M., Martindale, M. Q., Nardelli-Haeffliger, D., Baxter, E. & Price, D. J. Origin of segmental identity in the development of the leech nervous system. *Development* (Suppl. 2) 29–38 (1991).
- Wysocka-Diller, J. W., Aisemberg, G. O., Baumgarten, M., Levine, M. & Macagno, E. R. Characterization of a homologue of *bithorax*-complex genes in the leech *Hirudo medicinalis*. *Nature* **341**, 760–763 (1989).
- Dick, M. H. & Buss, L. W. A PCR-based survey of homeobox genes in *Ctenodrilus serratus* (Annelida: Polychaeta). *Mol. Phylogenet. Evol.* **3**, 146–158 (1994).
- Wong, V. Y., Aisemberg, G. O., Gan, W. B. & Macagno, E. R. The leech homeobox gene *Lox4* may determine segmental differentiation of identified neurons. *J. Neurosci.* **15**, 5551–5559 (1995).
- Kmita-Cunisse, M., Loosli, F., Biérne, J. & Gehring, W. J. Homeobox genes in the ribbonworm *Lineus sanguineus*: evolutionary implications. *Proc. Natl Acad. Sci. USA* **95**, 3030–3035 (1998).
- Snow, P. & Buss, L. W. *HOM/Hox* type homeoboxes from *Stylaria lacustris* (Annelida: Oligochaeta). *Mol. Phylogenet. Evol.* **3**, 360–364 (1994).
- Degnan, B. M. & Morse, D. E. Identification of eight homeobox-containing transcripts expressed during larval development and at metamorphosis in the gastropod mollusc *Helix pomatia*. *Mol. Mar. Biol. Biotechnol.* **2**, 1–9 (1993).
- Averof, M. & Akam, M. *HOM/HOX* genes in a crustacean: implications for the origin of insect and crustacean body plans. *Curr. Biol.* **3**, 73–78 (1993).
- Akam, M. *Hox* and *HOM*: homologous gene clusters in insects and vertebrates. *Cell* **57**, 347–349 (1989).
- Ruvkun, G. & Hobert, O. The taxonomy of developmental control in *Caenorhabditis elegans*. *Science* **282**, 2033–2041 (1998).
- Brooke, N. M., García-Fernández, J. & Holland, P. W. H. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* **392**, 920–922 (1998).
- Schubert, F. R., Nieselt-Struwe, K. & Gruss, P. The *Antennapedia*-type homeobox genes have evolved from three precursors separated early in metazoan evolution. *Proc. Natl Acad. Sci. USA* **90**, 143–147 (1993).
- Mackey, L. Y. *et al.* 18S rRNA suggests that Entoprocta are protostomes, unrelated to Ectoprocta. *J. Mol. Evol.* **42**, 552–559 (1996).
- Balavoine, G. Identification of members of several homeobox gene classes in a planarian using a ligation-mediated polymerase chain reaction technique. *Nucleic Acids Res.* **24**, 1547–1553 (1996).
- Swofford, D. L. *PAUP* Phylogenetic Analysis Using Parsimony (* and Other Methods)*, Version 4 (Sinauer, Sunderland, MA, 1998).
- Strimmer, K. & von Haeseler, A. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**, 964–969 (1996).
- García-Fernández, J. & Holland, P. W. H. Archetypal organization of the amphioxus *Hox* gene cluster. *Nature* **370**, 563–566 (1994).
- Holland, P. W. H. & García-Fernández, J. *Hox* genes and chordate evolution. *Dev. Biol.* **173**, 382–395 (1996).
- Duboule, D. *et al.* An update of mouse and human *HOX* gene nomenclature. *Genomics* **7**, 458–459 (1990).
- Wang, B. B. *et al.* A homeotic gene cluster patterns the anteroposterior body axis of *C. elegans*. *Cell* **74**, 29–42 (1993).
- Martinez, P., Rast, J. P., Arenas-Mena, C. & Davidson, E. H. Organization of an echinoderm *Hox* gene cluster. *Proc. Natl Acad. Sci. USA* **96**, 1469–1474 (1999).

Supplementary information is available on Nature's World-Wide Web site or on paper copy from the London editorial office of Nature.

Acknowledgements. We thank B. Winnepenninckx for *L. anatum* DNA; A. Knowlton, R. Highsmith, and P. Reynolds for *P. caudatus* tissue; E. Davidson for his interest and for sharing unpublished data; V. Kassner and N. Lartillot for their help; and A. Friday and G. Budd for comments on the manuscript. This work was supported by grants from the BBSRC and the Wellcome trust (M.E.A.); the Russian Basic Research Foundation and the Royal Society (T.A.); the CNRS, program 'Genome' and the Université Paris-Sud (A.A.); and the Howard Hughes Medical Institute (J.K.G. and S.B.C.).

Correspondence and requests for materials should be addressed to G.B. (e-mail: Guillaume.Balavoine@cgm.cnrs-gif.fr). Accession numbers: Lan: AF144672–AF144680; Nvi: AF151663–AF151673; Pca: AF144884–AF144893; Pvu: AF144666–144671.

A stop-codon mutation in the *BRI* gene associated with familial British dementia

Ruben Vidal*, Blas Frangione*, Agueda Rostagno*, Simon Mead†, Tamas Révész‡, Gordon Plant†‡ & Jorge Ghiso*

* Department of Pathology, New York University School of Medicine, 550 First Avenue, New York 10016, USA

† The National Hospital for Neurology and Neurosurgery, Queen Square, London WC1N 3BG, UK

‡ The Department of Neuropathology, Institute of Neurology, Queen Square, London WC1N 3BG, UK

Familial British dementia (FBD), previously designated familial cerebral amyloid angiopathy–British type¹, is an autosomal dominant disorder of undetermined origin characterized by progressive dementia, spasticity, and cerebellar ataxia, with onset at around the fifth decade of life. Cerebral amyloid angiopathy, non-neuritic and perivascular plaques and neurofibrillary tangles are the predominant pathological lesions^{1–4}. Here we report the identification of a unique 4K protein subunit named ABri from isolated amyloid fibrils. This highly insoluble peptide is a fragment of a putative type-II single-spanning transmembrane precursor that is encoded by a novel gene, *BRI*, located on chromosome 13. A single base substitution at the stop codon of this gene generates a longer open reading frame, resulting in a larger, 277-residue precursor. Release of the 34 carboxy-terminal amino acids from the mutated precursor generates the ABri amyloid subunit. The mutation creates a cutting site for the restriction enzyme *XbaI*, which is useful for detecting asymptomatic carriers. Antibodies against the amyloid or homologous synthetic peptides recognize both parenchymal and vascular lesions in FBD patients. A point mutation at the stop codon of *BRI* therefore results in the generation of the ABri peptide, which is deposited as amyloid fibrils causing neuronal disfunction and dementia.

The uninterrupted transmission from one generation to the next and analyses of segregation and sex ratios have previously confirmed the autosomal dominant mode of inheritance of FBD in a large family of more than 200 members encompassing seven generations¹. However, the biochemical basis of the disorder has remained elusive. It has been reported as an atypical form of familial Alzheimer's disease^{5–7}, an example of spongiform encephalopathies^{8–10} and a specific form of primary congophilic angiopathy¹¹. Classification attempts based on immunohistochemical analysis failed to demonstrate specific staining of the amyloid deposits with a large set of antibodies directed toward known amyloid molecules, although the lesions were immunoreactive for several amyloid-associated proteins (for example, apolipoproteins E and J and serum amyloid P-component¹²). C-terminal fragments of α - and β -tubulin have been reported to be associated with the amyloid lesions, although the identification of the major component of the amyloid deposits remained unknown¹³.

Using a combination of formic-acid solubilization and gel-filtration chromatography, we have isolated amyloid fibrils from case V41, a female member of this pedigree¹ who developed disease symptoms at age 56 years and died at age 65. Post-mortem examination revealed severe widespread cerebrovascular amyloidosis in the brain and spinal cord, non-neuritic amyloid plaques affecting the cerebellum, hippocampus, amygdala and occasionally the cerebral cortex, changes in the periventricular white matter, perivascular amyloid plaques and neurofibrillar degeneration in hippocampal neurons^{1,14}. A prominent component of relative molecular mass 4,000 (M_r 4K) was defined in SDS–polyacrylamide gels in preparations from leptomeningeal as well as parenchymal

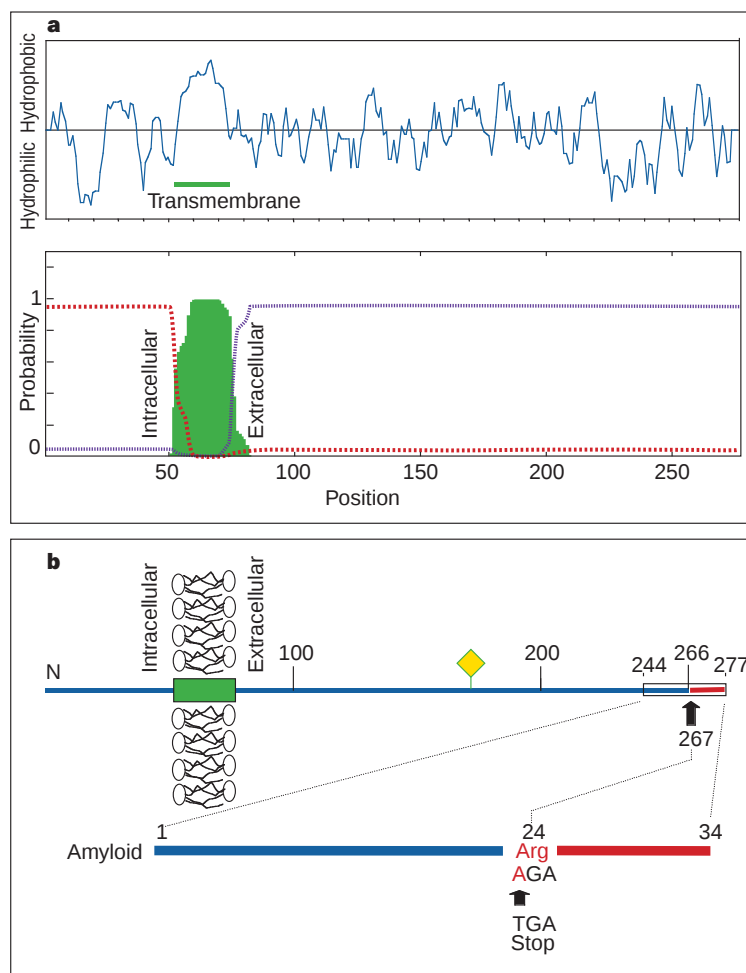


Figure 2 Model of the ABri precursor molecule as a cell-surface glycoprotein. **a**, Top panel, hydropathy plot of the putative ABri precursor according to ref. 17. Bottom panel, transmembrane prediction analysis according to ref. 18. **b**, Schematic representation of the ABri precursor molecule. Residues 52–74 comprise the putative single-transmembrane-spanning domain. The yellow diamond

denotes a single glycosylation site (position 170). The box encompassing positions 244–277 indicates the location of the 34-amino-acid ABri amyloid in the mutated precursor protein. The arrow points to the mutated codon (at position 267) resulting in the change of a stop codon TGA into an arginine residue (AGA).

170. Neither a signal sequence for transport through the membrane of the endoplasmic reticulum nor distinct motifs of known protein families could be detected. Homology searches done with the sequences retrieved for the human cDNA showed that ESTs of chicken, rat, mouse, rabbit and pig origin had been deposited in the databanks, their ORFs being highly homologous to the human sequence (for example, identity between human and murine is 96%). Studies done with mouse homologues indicate that the protein may belong to a multigene family¹⁹.

The gene encoding the BRI transcript was mapped to human chromosome 13 by fluorescence in-situ hybridization (FISH) analysis. The initial FISH experiment resulted in specific labelling of the long arm (q) of a group D chromosome, which was believed to be chromosome 13 on the basis of size, morphology and banding pattern (Fig. 3a). A second experiment was conducted in which a biotin-labelled probe that was specific for the centromere of chromosomes 13 and 21 was co-hybridized with the BRI clone, which resulted in the specific labelling of the centromere (red) and the long arm (green) of chromosome 13 (Fig. 3b). Measurements of ten specifically labelled chromosomes 13 showed that the BRI gene is located at a position that is 34% of the distance from the centromere to the telomere of chromosome 13q, an area that corresponds to band 13q14 (Fig. 3c). A search of sequence tagged sites (STS) databases (GenBank, NCBI) confirmed the results described above (accession numbers G25377 and G28192).

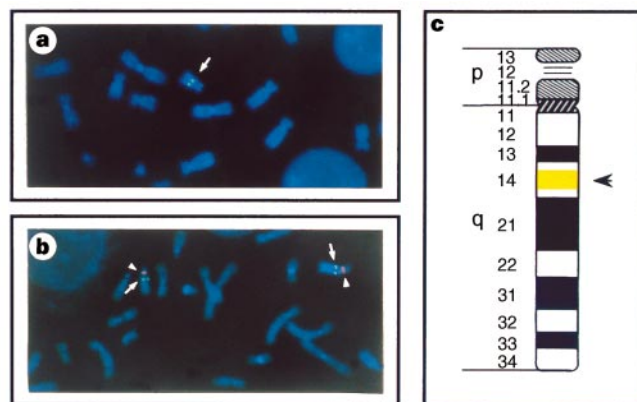


Figure 3 Chromosomal localization of *BRI*. **a**, One-colour experiment. Digoxigenin-labelled *BRI*-specific genomic probe was detected with fluoresceinated antidigoxigenin (arrow). **b**, Two-colour experiment. Biotin-labelled probe specific for the centromere of chromosomes 13 and 21 was detected by Texas red avidin (arrowhead), whereas digoxigenin-labelled *BRI* was detected as above (arrow). A total of 80 metaphase cells were analysed with 72 exhibiting specific labeling. **c**, Ideogram illustrating the chromosomal position of *BRI* at band 14 on the long arm of chromosome 13 (13q14). (Ideogram reproduced from the International System for Human Cytogenetic Nomenclature, 1995.)

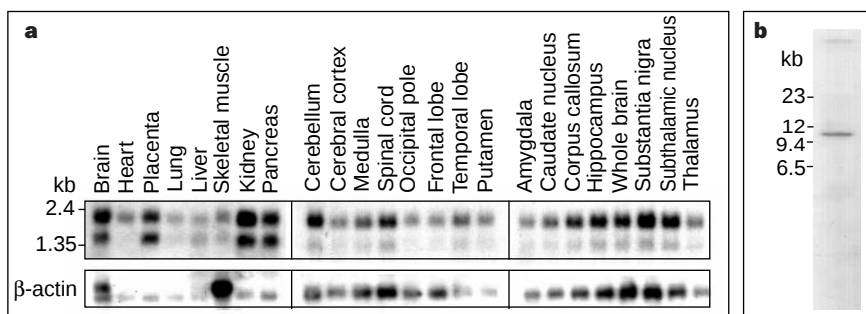


Figure 4 Expression of the ABri precursor protein. **a**, Northern blots identify two mRNA transcripts of ~2.0 and 1.6 kb in different tissues and various regions of the human brain. Controls with β -actin are shown in the bottom panel. **b**, Southern

hybridization. A single ~10.5-kb fragment of *Bam*HI-digested genomic DNA hybridized with the p701 BRI cDNA probe. Positions of the *Hind*III-digested λ DNA and 1-kb DNA ladder (Gibco BRL) are marked on the left side.

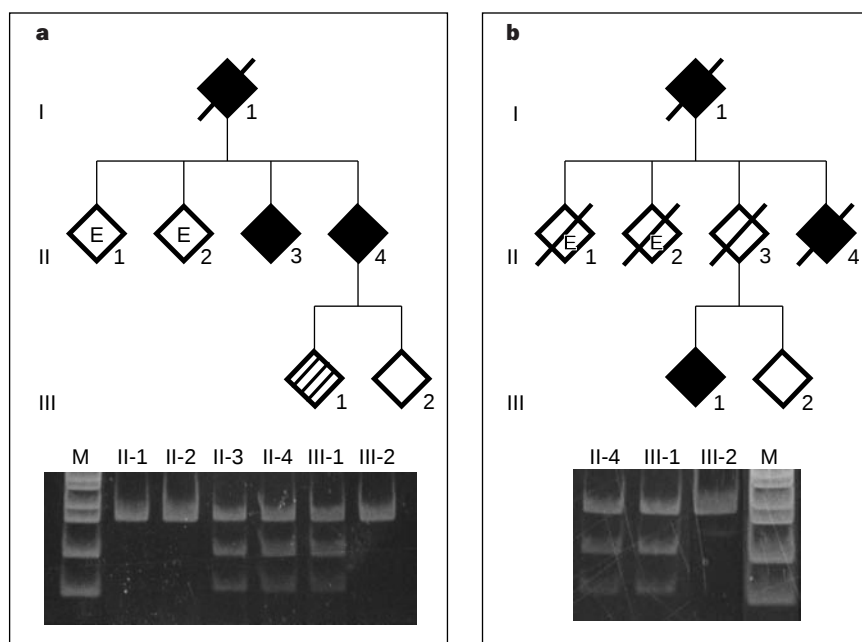


Figure 5 Analysis of nucleotide substitutions in the gene encoding the ABri precursor protein. Members of the British family pedigree are shown in **a** and **b**. Roman numbers indicate different generations. Individual IA(1) is a descendant of affected case III(1), and individual IB(1) is a descendant of affected case III(3) as described¹. M indicates ϕ X174 RF DNA *Hae*III fragments. In wild-type sequences,

no cleavage is observed, whereas in mutant sequences the PCR product is cleaved into 75- and 116-bp fragments. Affected members (filled symbols) and at-risk members (hatched symbols) carried one allele with the *Xba*I site. The *Xba*I site was not present in all obligate escapees (open symbol with E) and all normal controls (not shown).

Hybridization of the clone of the human precursor protein to northern blots identified a major messenger RNA transcript of ~2.0 kilobases (kb) and a second transcript of ~1.6 kb, which were expressed in most regions of the human brain as well as in several peripheral tissues (Fig. 4a). High expression was observed in the brain, placenta, kidney and pancreas, whereas lower expression levels were observed in the heart, lung, liver and skeletal muscle. Northern analysis of different brain regions detected predominantly the 2.0-kb transcript in the cerebellum, spinal cord, subthalamic nucleus, substantia nigra and hippocampus, whereas lower expression levels were observed in the cerebral cortex, amygdala and thalamus. All the retrieved cDNA clones that include both 5' and 3' untranslated regions (UTRs) have a size in the range of the two signals observed in the northern hybridizations. Because a single-specific DNA fragment was observed in Southern blots using the same probe (Fig. 4b), the two mRNA transcripts may represent either alternatively spliced or polyadenylated isoforms. As indicated in Fig. 1a, the 3' region of the precursor protein contains five polyadenylation signals. Transcripts using the first signal, ATTAAA, were found in ESTs from placenta (AI189911, AI200443 and R76391), fetal liver/spleen (AI248856, AI023889 and H73884), aorta (D58336, D57790, D57728 and D57976), white blood cells

(AH356254), parathyroid tumours (AI075223 and AI168126), pregnant uterus (AA135881) and fetal heart (W67150). Transcripts that used the second polyadenylation signal, TATAAA, were found in breast (R86131), pregnant uterus (AA149499), pancreas (AA113057), cerebellum (AA323476 and AA325464), parathyroid tumour (W56225), placenta (N40489) and heart (C04137). Two perfect polyadenylation signals, AATAAA, are followed by a fifth poly-A signal, ATTAAA, all reported in various ESTs (THC212690) derived from placenta, liver/spleen, heart, parathyroid tumours, pregnant uterus, pancreas, cerebellum, fibroblasts, white blood cells, colon, testis, lung, epididymus, kidney, fetal heart, fetal liver/spleen, trabecular bone cells, senescent fibroblasts and infant brain.

Nucleotide sequence analysis of the protein transcript isolated from seven affected members of the FBD family revealed a single nucleotide transition (T for A) at stop codon 267, resulting in the presence of an arginine residue and an ORF that was 33 nucleotides longer and coded for a precursor of 277 amino acids (instead of 266) (Figs 1b, 2b). The mutation was not present in non-affected individuals ($n = 7$) of the same kindred. The change of a stop codon TGA into an arginine (AGA) introduced a cutting site for the restriction enzyme *Xba*I, which was detected by digestion of polymerase chain reaction (PCR)-amplified genomic DNA with this

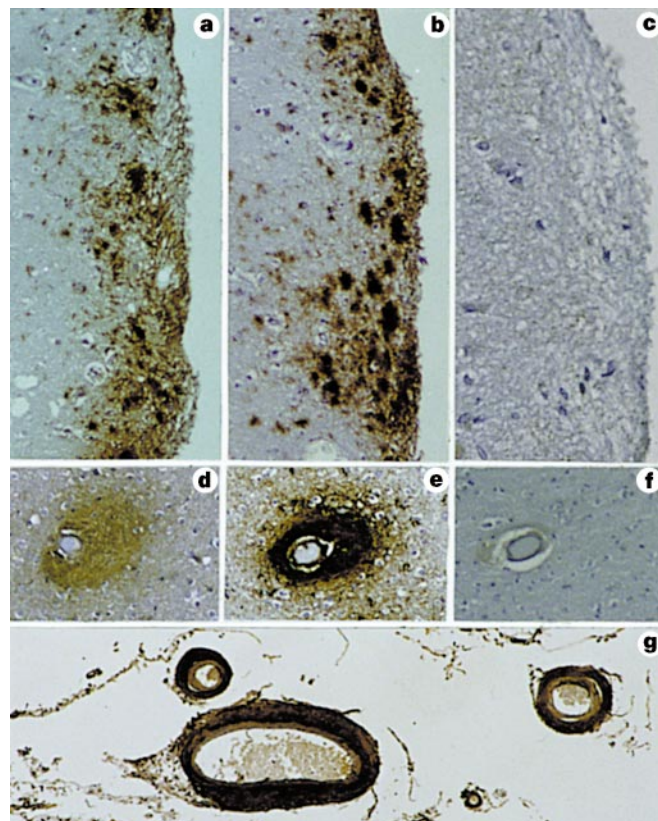


Figure 6 Immunohistochemical analysis. Staining of temporal adjacent sections of case V41 (ref. 1) with antibodies 547 (anti-ABri amyloid; 1:300) and 338 (anti-C-terminal synthetic peptide; 1:1,000). **a, b**, Parenchymal non-neuritic plaques immunoreactive with Abs 547 and 338. **d, e**, Perivascular plaques specifically labelled with Abs 547 and 338. **g**, Leptomeningeal vessels immunostained with Ab 338. No immunoreactivity was seen after absorption of Ab 547 (not shown) and 338 (**c, f**). Original magnifications, $\times 250$ (**a–f**) and $\times 40$ (**g**) (paraffin sections, haematoxylin counterstain).

restriction endonuclease. The amplification product (191 bp) remains undegraded in wild-type sequences, whereas in mutant sequences the PCR product is cleaved into two fragments of 75 bp and 116 bp (Fig. 5). Affected members and at-risk members tested so far proved to be heterozygous, carrying only one allele with the *Xba*I site (some examples are shown in Fig. 5). These results were confirmed by DNA sequencing after cloning of the amplified fragments into pCR2.1 vector (Invitrogen). The nucleotide substitution is associated with the disease, and was not seen in asymptomatic family members, individuals with unrelated neurologic disorders ($n = 42$) or normal controls ($n = 71$) from comparable ethnic origins. Furthermore, the stop codon that is altered in the British family is conserved in the murine homologue, indicating that the nucleotide substitution is a pathogenic mutation rather than innocent polymorphism. Examples of stop-codon mutations are abundant in the literature, one of the most common being the creation of a stop codon that produces a shorter protein. The abolishment of a stop codon and the appearance of a longer ORF that codes for a sequence extending to the next in-frame stop codon is unusual. The best known examples are the mutants in the globin moiety of haemoglobin, all of them related to thalassaemia.

The purified amyloid ($M_r = 4K$) as well as a synthetic peptide comprising the last ten residues of the ABri sequence were used to raise polyclonal antibodies in rabbits. Both antibodies specifically recognized the amyloid deposits (Fig. 6). Anti-ABri amyloid (antibody 547; Fig. 6a, d) recognized the same type of lesions as were immunodetected by the antisynthetic peptide C-terminal ABri (antibody 338; Fig. 6b, e) although the latter showed stronger immunoreactivity. The specificity of the immunostaining was

corroborated by absorption of both antibodies with a synthetic peptide that was homologous to the 34-amino-acid, full-length ABri amyloid (Fig. 6c, f). The immunoreactivity co-localized with yellow–green birefringent material observed under polarized light after Congo red staining (not shown). The presence of non-neuritic and perivascular plaques was revealed by immunoreactivity with the antibodies. Leptomeningeal vessels full of amyloid were clearly detected by the antibody 338 (Fig. 6g). Identical results were obtained with antibody 547. No immunoreactivity using 547 and 338 antibodies was observed in brain sections of sporadic cerebral amyloid angiopathy (CAA), sporadic Alzheimer's disease, Down's syndrome, hereditary cerebral haemorrhage with amyloidosis-Dutch type, hereditary cerebral haemorrhage with amyloidosis-Icelandic type, Hungarian transthyretin cerebral amyloidosis, and systemic cases of light-chain amyloidosis (kidney), light-chain deposition disease (kidney) and amyloid A (heart)^{20,21}.

FBD is similar to Alzheimer's disease, particularly to the PS1- $\Delta 9$ variant described recently²². Both disorders have in common the production of peptides of $M_r \sim 4K$ (ABri and amyloid- β or A β), originated by the processing of different cell-membrane precursor proteins encoded by genes located on chromosome 13 and 21, respectively. In addition, the cytoskeletal pathology of FBD is undistinguishable from that observed in other neurodegenerative diseases, including sporadic and familial forms of Alzheimer's disease and the vascular form of prion protein (PrP)-CAA²³. Our study supports the proposal that amyloid proteins are important in the initiation of neuronal dysfunction. Understanding the pathogenesis of the rare disease FBD may shed light on the cause of dementia in Alzheimer's disease as well as in other neurological disorders related to alterations in protein folding and/or aggregation. □

Methods

Isolation and characterization of ABri amyloid protein. Amyloid was isolated from leptomeninges and parenchymal non-neuritic plaques of case V41 (ref. 1) using a combination of phosphate buffered saline (PBS) extraction, collagenase digestion, formic acid solubilization and gel-filtration chromatography. The resulting subunit ($M_r = 4K$) was subjected to N-terminal sequence analysis on a 477A protein sequencer with an on-line 120A PTH analyser (Applied Biosystems). Internal sequence information was obtained from tryptic peptides purified by HPLC. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) was carried out on an α -cyano-4-hydroxy-cinnamic acid matrix using the dried droplet method. The search of ESTs databanks was done with the aid of the BLAST algorithm (<http://www.ncbi.nlm.nih.gov/BLAST>). Hydrophobicity indices from ref. 17 (using a window of seven residues) were calculated with GPMW Ver 3.10 software (Lighthouse data, Denmark). Transmembrane prediction analysis¹⁸ was done at the web site of the Prediction Servers, Center for Biological Sequence Analysis (<http://www.cbs.dtu.dk>).

cDNA cloning. The complete nucleotide sequence of a human transcript of ~ 1.8 kb was obtained from the consensus of independent cDNA clones recovered by standard 5' and 3' rapid amplification of cloned ends (RACE) using Marathon-ready cDNA from kidney and brain cDNA libraries (Clontech), and the sequence of EST clones AA149631, AA149499, R76391 and R86131 and clone THC212690 (TIGR gene index database; <http://www.tigr.org>). EST clones AA149631 and AA149499 were obtained from ATCC (Rockville) and sequenced completely in both directions by automated cycle sequencing (ABI).

Fluorescence in situ hybridization. A P1 human genomic DNA clone containing the *BRI* gene (~ 75 kb) was labelled with digoxigenin dUTP by nick translation. Labelled probe was combined with sheared DNA and hybridized to normal metaphase chromosomes derived from phytahaemagglutinin (PHA)-stimulated peripheral blood lymphocytes in a solution containing 50% formamide, 10% dextran sulphate and $2 \times$ SSC. Specific hybridization signals were detected by incubating the hybridized slides in fluoresceinated antidigoxigenin antibodies followed by counter staining with 4',6-diamidindine-2-phenylindole (DAPI, Boehringer) for the one-colour experiment. Probe detection for the two-colour experiment was accomplished by incubating the slides in fluoresceinated antidigoxigenin antibodies and Texas red avidin followed by counterstaining with DAPI.

Northern blot analysis. A cDNA probe spanning BRI nucleotides 268 to 969 (named p701) was hybridized to northern blots (Clontech) containing 2 µg per lane of human poly-A⁺ mRNA isolated from different tissues and various regions of the human brain. Filters were prehybridized at 42 °C for 3 h in 5 × SSC, 5 × Denhardts, 20 mM NaH₂PO₄, 0.2% SDS, 20% formamide and 250 µg ml⁻¹ salmon sperm DNA. Hybridization was performed for 16 h at 42 °C in the same solution containing the ³²P-labelled p701 probe at 1 × 10⁶ c.p.m. per ml. Membranes were washed at 42 °C in 2 × SSPE containing 0.1% SDS. Blots were rehybridized in ExpressHyb solution (Clontech) at 68 °C for 1 h using the ³²P-labelled β-actin probe.

Southern blot analysis. Ten micrograms of genomic DNA for Southern blot analysis (Clontech) was digested with *Bam*HI, size-fractionated on 1% agarose gel, blotted onto a nitrocellulose filter and hybridized to ³²P-labelled p701 probe at 42 °C. After post-hybridization washes at 65 °C in 2 × SSC, 0.1% SDS, the blots were exposed to X-ray film at -80 °C.

Restriction analysis. Amplification of genomic DNA samples isolated from peripheral blood leukocytes of living family members and autopsy tissue was done using oligonucleotide primer pairs F (5'-CGTGAAGCCAGCAATT GTTTCGCA-3') and R (5'-AGCCCTGTTTGCTACTTACATG-3') (the product was 191 bp) by PCR using 250 µmol dNTPs, 2.5 mM MgCl₂, 50 pmol oligonucleotides, in 100 µl, and cycled for 30 cycles of 94 °C for 30 s, 50 °C for 30 s and 72 °C for 30 s. PCR products were cloned and sequenced by automated cycle sequencing. Analysis of the amplified PCR products was done with the restriction enzyme *Xba*I (Gibco, BRL) according to the manufacturer's instructions, and the resulting products were resolved by non-denaturing 5% polyacrylamide gel electrophoresis.

Immunohistochemistry. The 4K purified ABri amyloid as well as a synthetic peptide CTVKKNIEEN, which is homologous to the ten C-terminal residues of the ABri molecule and which contains an extra N-terminal cysteine for coupling to keyhole limpet haemocyanin, were used to induce a polyclonal antibody response in rabbits. After an initial stimulation with 200 µg of antigen emulsified in monophospholipid A-synthetic trehalose dicorynomycolate (RIBI; Immunochem Research) adjuvant, animals were boosted with 50 µg of antigen every three weeks for twelve weeks. Specific antibodies were tested by enzyme-linked immunosorbent assay (ELISA) and dot-blot analysis against a synthetic peptide that was homologous to the full-length ABri sequence. Immunoglobulin-γ (IgG) fractions were purified from the rabbit serum by means of specific binding to protein G (Gammabind G, Pharmacia). Immunoabsorption of both antibodies was done with full-length ABri peptide (100 µg of peptide per 50 µl of antibody) for 1 h at 37 °C and 16 h at 4 °C, followed by centrifugation at 14,000g for 5 min. Temporal adjacent sections of case V41 (ref. 1) were immunostained with both antibodies followed by biotinylated anti-rabbit IgG and streptavidin-horseradish peroxidase (HRP). Colour was developed with diaminobenzidine tetrahydrochloride (DAB) and hydrogen peroxide.

Received 15 February; accepted 4 May 1996.

- Plant, G. T., Révész, T., Barnard, R. O., Harding, A. E. & Gautier-Smith, P. C. Familial cerebral amyloid angiopathy with nonneurotic plaque formation. *Brain* **113**, 721-747 (1990).
- Worster-Drought, C., Hill, T. R. & McMenemey, W. H. Familial presenile dementia with spastic paralysis. *J. Neurol. Psychopathol.* **14**, 27-34 (1933).
- Worster-Drought, C., Greenfield, J. G. & McMenemey, W. H. A form of familial presenile dementia with spastic paralysis (including the pathological examination of a case). *Brain* **63**, 237-254 (1940).
- Worster-Drought, C., Greenfield, J. G. & McMenemey, W. H. A form of presenile dementia with spastic paralysis. *Brain* **67**, 38-43 (1944).
- Corsellis, J. & Brierley, J. B. An unusual type of presenile dementia: (atypical Alzheimer's disease with amyloid vascular change). *Brain* **77**, 571-587 (1954).
- Aikawa, H., Suzuki, K., Iwasaki, Y. & Iizuka, R. Atypical Alzheimer's disease with spastic paresis and ataxia. *Ann. Neurol.* **17**, 297-300 (1985).
- Masters, C., Gajdusek, C. & Gibbs, C. J. The familial occurrence of Creutzfeldt-Jakob disease and Alzheimer's disease. *Brain* **104**, 535-558 (1981).
- Keohane, C., Peatfield, R. & Duchon, L. W. Subacute spongiform encephalopathy (Creutzfeldt-Jakob disease) with amyloid angiopathy. *J. Neurol. Neurosurg. Psych.* **48**, 1175-1178 (1985).
- Courten-Myers, G. & Mandybur, T. I. Atypical Gerstmann-Sträussler syndrome or familial spinocerebellar ataxia and Alzheimer's disease? *Neurology* **37**, 269-275 (1987).
- Pearlman, R. L., Towfigh, J., Pezeshkpour, G. H., Tenser, R. B. & Turel, A. P. Clinical significance of types of cerebellar amyloid plaques in human spongiform encephalopathies. *Neurology* **38**, 1249-1254 (1988).
- Vinters, H. Cerebral amyloid angiopathy: a critical review. *Stroke* **18**, 311-324 (1987).
- Ghiso, J., Plant, G. T., Révész, T., Wisniewski, T. & Frangione, B. Familial cerebral amyloid angiopathy (British type) with nonneurotic amyloid plaque formation may be due to a novel amyloid protein. *J. Neurol. Sci.* **129**, 74-75 (1995).
- Baumann, M. H., Wisniewski, T., Levy, E., Plant, G. T. & Ghiso, J. C-terminal fragments of α- and β-tubulin form amyloid fibrils *in vitro* and associate with amyloid deposits of familial amyloid angiopathy, British type. *Biochem. Biophys. Res. Commun.* **219**, 238-242 (1996).
- Révész, T. et al. Cytoskeletal pathology in familial amyloid angiopathy (British type) with non-neurotic plaque formation. *Acta Neuropath.* **97**, 170-176 (1999).

- Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402 (1997).
- Kozak, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNA. *Nucleic Acids Res.* **15**, 8125-8148 (1987).
- Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132 (1982).
- Sonnhammer, E. L., von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *LSMB.* **6**, 175-182 (1998).
- Deleersnijder, W. et al. Isolation of markers for chondro-osteogenic differentiation using cDNA library subtraction. *J. Biol. Chem.* **271**, 19475-19482 (1996).
- Ghiso, J., Wisniewski, T. & Frangione, B. Unifying features of systemic and cerebral amyloidosis. *Mol. Neurobiol.* **8**, 49-64 (1994).
- Vidal, R. et al. Meningocerebrovascular amyloidosis associated with a novel transthyretin mis-sense mutation at codon 18 (TTRD18G). *Am. J. Pathol.* **148**, 361-366 (1996).
- Crook, R. et al. A variant of Alzheimer's disease with spastic paraparesis and unusual plaques due to deletion of exon 9 of presenilin 1. *Nature Med.* **4**, 452-455 (1998).
- Ghetti, B. et al. Vascular variant of prion protein cerebral amyloidosis with tau-positive neurofibrillary tangles: the phenotype of the stop codon 145 mutation in PRNP. *Proc. Natl Acad. Sci. USA* **93**, 744-748 (1996).

Acknowledgements. This work was supported by the NIA LEAD award Alzheimer's disease and amyloid proteins (B.F.). J.G. is the recipient of the NIDA from AHA (NYC affiliate).

Correspondence and requests for materials should be addressed to J.G. (e-mail: ghisoj01@popmail.med.nyu.edu).

Human theta oscillations exhibit task dependence during virtual maze navigation

Michael J. Kahana^{*†}, Robert Sekuler^{*†}, Jeremy B. Caplan^{*}, Matthew Kirschen^{*} & Joseph R. Madsen^{*†‡}

^{*} Volen Center for Complex Systems, Brandeis University, Waltham, Massachusetts 02454, USA

[†] Department of Neurosurgery, Children's Hospital, Boston, Massachusetts 02115, USA

[‡] Department of Surgery, Harvard Medical School, Boston, Massachusetts 02115, USA

Theta oscillations (electroencephalographic activity with a frequency of 4-8 Hz) have long been implicated in spatial navigation in rodents¹⁻³; however, the role of theta oscillators in human spatial navigation has not been explored. Here we describe subdural recordings from epileptic patients learning to navigate computer-generated mazes. Visual inspection of the raw intracranial signal revealed striking episodes of high-amplitude slow-wave oscillations at a number of areas of the cortex, including temporal cortex. Spectral analysis showed that these oscillations were in the theta band. These episodes of theta activity, which typically last several cycles, are dependent on task characteristics. Theta oscillations occur more frequently in more complex mazes; they are also more frequent during recall trials than during learning trials.

Theta oscillations are seen when animals engage in a variety of spatial, motor and cognitive tasks¹⁻⁹. For example, when rodents explore environments, theta oscillations are specifically related to the firing of hippocampal place cells that encode the animals' position^{2,3}. Such findings have led researchers to assign a crucial functional role to the theta activity in physiologically based models of behaviour¹⁰⁻¹². Although most studies of theta waves in rodents focus on the hippocampus and the role of the medial septum, they have also been seen in other brain regions, including the hypothalamus, entorhinal cortex, cingulate cortex and superior colliculus¹³⁻¹⁵. It is unclear whether cortical theta activity results from cortico-hippocampal interactions or the presence of multiple theta generators. There has been uncertainty over the existence (and possible functional role) of task-dependent theta activity in humans¹⁶⁻¹⁸. To help resolve this uncertainty we devised a spatial learning task for humans that resembles those tasks that elicit strong theta oscillations in rodents.

Although neuroimaging studies have examined brain activity during spatial navigation and wayfinding tasks¹⁹⁻²³, the techniques

used (positron emission tomography and functional magnetic resonance imaging) lack the temporal resolution needed to observe spectrally distinct brain activity such as theta oscillations. We have recorded intracranially from multiple cortical loci in humans who are exploring an unfamiliar, computer-simulated environment. Our results demonstrate both the existence and task dependence of theta oscillations during maze learning.

We tested three patients with medically intractable epilepsy who had arrays of subdural electrodes implanted for one to two weeks to localize the site or sites of seizure onset. The placement of these arrays was determined by the clinical team so as to best localize suspected epileptogenic foci and to identify functional regions to be avoided in surgery.

Unlike scalp electroencephalograms (EEG), intracranial EEG (iEEG) provides direct access to population activity in theoretically important cortical regions located on the brain's ventral surface. In addition, iEEG offers improved spatial resolution without attenuation of brain signals by the skull and scalp. iEEG also avoids muscle artefacts, including eye movements²⁴.

Subjects used a computer keyboard to navigate through visually rich, computer-rendered, three-dimensional mazes. Each virtual maze comprised a series of corridors, each leading to a T-junction (Fig. 1a). The subject's aim was to navigate as quickly and accurately as possible from a starting point to a goal position.

Subjects traversed each maze in two modes, denoted Study and Test. In Study mode, arrows placed at junctions directed subjects through the maze (Fig. 1b). In Test mode, the arrows were removed, forcing subjects to rely on a learned representation of the maze. Subjects navigated each maze four times in Study mode, and then navigated the same maze in Test mode until they could traverse the maze three times successively without errors.

Figure 2a shows a sample iEEG trace recorded from an electrode on the inferior frontal gyrus in subject 1 (Talairach²⁵ coordinates: L-R = -49, A-P = +21, I-S = -4 mm). This trace represents brain activity during a single maze traversal in Test mode. The complex waveform is interrupted by episodes of rhythmic slow-

wave activity (Fig. 2a). These slow waves are in the theta range and appear during both Study and Test trials in all three subjects, and in many brain regions.

Our finding that theta activity is not continuous but occurs in distinct, well defined episodes can be seen in the frequency-time spectrogram (Fig. 2b) for each maze. Z-scores, thresholded at $z = 1$, are plotted for task power relative to the distribution of power across all maze trials (both long and short). The overall peak in the theta range can be seen consistently across both Study and Test trials (Fig. 2c).

To determine how theta episodes were related to specific attributes of our maze-learning task we compared power spectra obtained during long (12-junction) and short (6-junction) mazes. We had previously established that 12-junction mazes pose a serious challenge for any inexperienced subject, whereas 6-junction mazes are typically mastered by the first test trial (data not shown). Consistent with these normative findings, our three subjects took approximately 10 trials to learn long mazes, but only one or two learning trials to master short mazes. To test whether theta activity was sensitive to the increased cognitive demands of the long mazes we compared the probability of a theta episode occurring during performance of long versus short mazes. For each subject, at certain cortical loci, the percentage of maze trial time spent in theta episodes (P_θ) was greater for long mazes than for short mazes ($P < 0.01$ by a two-tailed Mann-Whitney U -test). In those electrodes showing significantly higher P_θ for long mazes as compared with short mazes, mean P_θ was 15% for long mazes and 10% for short mazes.

The median duration of theta episodes in long mazes was 720 ms (95th percentile = 1,630 ms) and in short mazes the median duration was 685 ms (95th percentile = 1,405 ms). These values were computed only across electrodes that showed a significant effect of maze complexity. We found no electrodes that recorded more theta

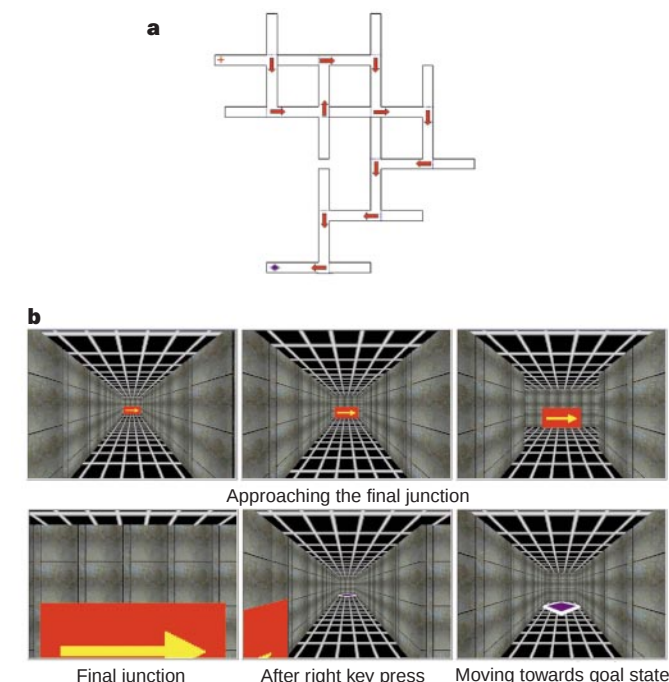


Figure 1 Navigation of computer-rendered multiple T-junction mazes. **a**, Blueprint of a sample 12-junction maze. Plus sign marks start position; diamond marks finish position. Arrows indicate the correct path. Dotted lines mark invisible barriers. **b**, Six sample views of the virtual environment, such as a subject would see when traversing a maze.

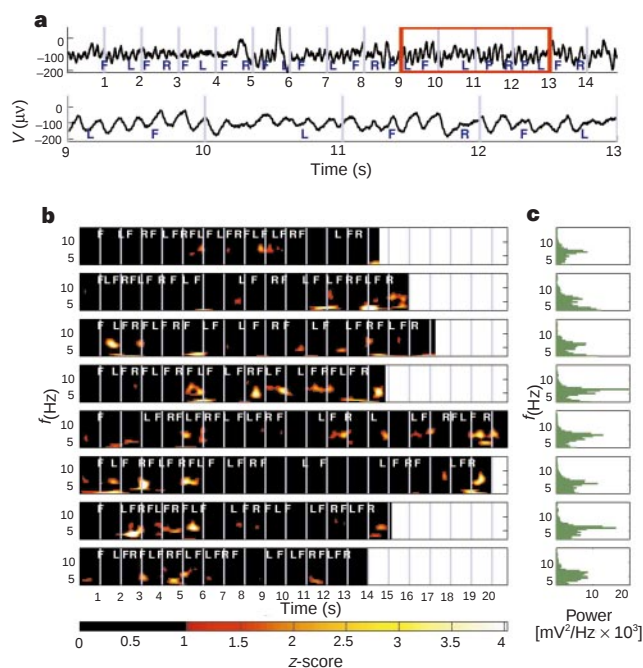


Figure 2 Theta oscillations revealed by intracranial EEG. **a**, Upper trace, sample iEEG trace from study trial 4, maze 15 recorded from an electrode on the inferior frontal gyrus in subject 1. Key presses are indicated by F (forward), R (right) and L (left). Lower trace, enlarged view of the boxed region, emphasizing the theta oscillation. **b**, Time-frequency spectrograms for all trials in maze 15 recorded from the same electrode as in **a**. The colour scale (bottom) reflects Z-score relative to wavelet-window power across all maze trials. Rows 1–4, study trials; row 5, a test trial in which the subject made an error; rows 6–8, error-free test trials. **c**, Average power for each trial.

activity for short mazes than for long mazes. The locations of the 48 (out of 171) electrodes that exhibited this task complexity effect are shown in Fig. 3a–c. This difference in P_θ also showed up as a difference in the average power spectra for most of the significant electrodes, with the same direction of difference between long and short maze trials (Fig. 3d–g).

Research on learning has been advanced by the theoretical distinction between encoding and retrieval processes²⁶. To investigate whether theta activity might have a different role in retrieval from its role in encoding we examined theta episodes separately in the first four Study trials and the final three Test trials across all mazes. In 45 electrodes, theta episodes occurred more frequently during Test trials than during Study trials ($P < 0.01$). No electrodes showed the reverse pattern.

As has been found in rodents, theta oscillations in our experiments with human subjects occur during spatial navigation. These oscillations are visually striking against a complex background signal. However, our recordings are from the cortical surface and not from the hippocampus itself. Further work is required before we can make precise statements about the link between our observation of human theta waves and the literature on hippocampal and cortical theta activity in rodents.

Because our observations came only from individuals with epilepsy, we cannot guarantee that non-epileptic brains would show theta oscillations identical to those reported here. To reduce the sensitivity of our analysis to epileptic artefacts, we focused not on the presence of theta activity *per se*, but on the task-related properties of theta episodes. Additionally our results are not artefacts of the intermittent sharp transients (70–200 ms) sometimes seen in epileptic brains. Indeed, our results are essentially unchanged when the criterion for a theta episode is decreased from 500 to 167 ms. Further, although in subjects 2 and 3 there is some overlap between the epileptogenic tissue and the electrodes showing task-related theta activity, in subject 1, whose raw signal showing theta oscillations is illustrated in Fig. 2, electrodes showing task-related theta were not near the clinically identified epileptogenic focus.

In summary, our spatial learning task not only reveals clear

evidence of theta activity, but also shows that it occurs in distinct, task-dependent episodes. Because these theta episodes become more frequent in long mazes and during retrieval, it is likely that theta oscillations are important in human spatial cognition. □

Methods

Subjects. Our three subjects were of average range personality and intelligence and were all able to learn the mazes. Subject 1 (female, age 16) and subject 3 (male, age 19) performed within the range of college students, while subject 2 (male, age 14) had more difficulty. Our research protocol was approved by the institutional review board at Children's Hospital and we obtained informed consent from subjects and their guardians.

Procedure. Subjects learned to navigate computer-rendered multiple-T-junction mazes. Because their visual features were uniform, the mazes offered subjects no information about their position. The subjects' position and viewing angle were manipulated to simulate virtual movement. A subject could move in three different ways: pressing the forward/up arrow key moved the subject down a corridor over the course of 280 ms, and pressing the left or right arrow instantly changed the subject's view by 90 degrees in the respective direction. The maze had to be navigated completely from start to finish. Although these constraints compromised some of the freedom of the subjects to move *ad lib* through a maze, they equated the data format across error-free trials. The sequence of successive turns excluded maze sequences with three consecutive turns in the same direction, preventing the maze from crossing itself. We created a random pool of 6-, 8- and 12-junction mazes. Invisible barriers kept subjects from moving down incorrect corridors.

One testing session involved navigating 20 different mazes. The first four mazes were practice mazes (8-junction), excluded from analysis. Eight mazes were short (6-junction) and eight were long (12-junction), presented in random order. Subjects 2 and 3 each performed one session and subject 1 performed two sessions.

Intracranial EEG recording. Grids and strips contained multiple platinum electrodes with inter-electrode spacing of 1 cm. The locations of the electrodes were determined using co-registered postoperative computed tomograms and preoperative magnetic resonance imaging by an indirect stereotactic technique²⁵. The iEEG signal was amplified, band-pass filtered (0.5–100 Hz single pole), digitized at 200 Hz, recorded on analogue videotape using Telefactor apparatus and then copied from videotape to disk. In this conversion process, degradation of the recording medium caused the loss of occasional samples. Missing samples were assigned interpolated values equal to the average of the two adjacent sample values.

Spectral analysis. Spectrograms (Fig. 2b) were computed using a Morlet wavelet transform²⁷ with a sliding window (16 cycles) in 0.5 Hz increments. spectrogram values were thresholded at mean + one standard deviation ($z = 1$) computed from all maze trials across all wavelet windows. Average power spectra were computed in MATLABTM with a Welch's window (window, 2 s; overlap, 1 s). Trial power (Fig. 2c) was computed across all fast Fourier-transform windows for the trial. Power spectra for a given maze length (Fig. 3d–g) were computed across all windows in all trials of mazes with a given length. A theta episode is defined as a duration longer than 500 ms (roughly three theta cycles) throughout which the power exceeds $z = 1$ (relative to the distribution of wavelet-window power across all maze trials) on at least one frequency in the range 4–8 Hz (1 Hz increments). For a given trial, the percentage of maze trial time spent in theta episodes (P_θ) is the total time in theta episodes divided by the total time in the trial.

Resections. All subjects underwent subsequent resection of epileptogenic tissue (subject 1, right frontal lobe focus 2 cm × 3 cm, involving orbital gyrus; subject 2, right inferior occipital and inferomedial temporal cortex from a posterior approach; subject 3, standard anterior left temporal lobectomy, including amygdala and some hippocampus).

Received 5 February; accepted 12 April 1999.

1. Winson, J. Loss of hippocampal theta rhythm results in spatial memory deficit in the rat. *Science* **201**, 160–163 (1978).
2. O'Keefe, J. & Recce, M. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* **3**, 317–330 (1993).
3. Skaggs, W. E., McNaughton, B. L., Wilson, M. A. & Barnes, C. A. Theta phase precision in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* **6**, 149–172 (1996).

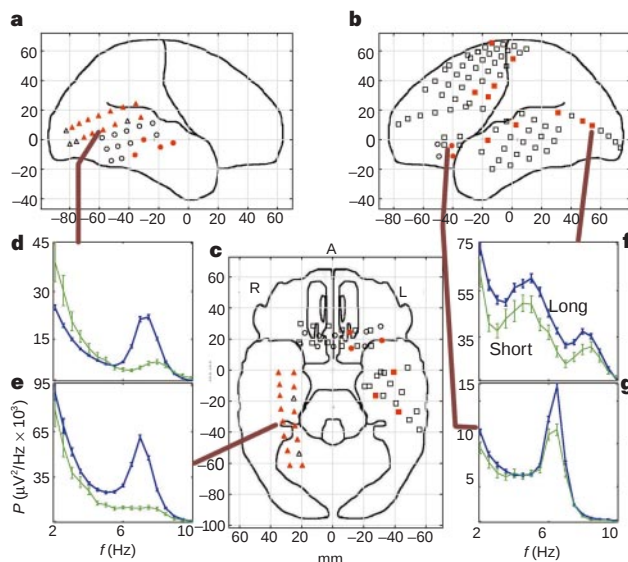


Figure 3 Electrode locations and task-dependent theta activity seen for all three subjects, shown on three diagrams of a standard brain. **a**, Right lateral view. **b**, Left lateral view. **c**, Inferior view. Circles, triangles and squares represent electrodes in subjects 1, 2 and 3, respectively. Electrodes filled in red recorded increased theta activity during long mazes relative to short mazes ($P < 0.01$, two-tailed Mann-Whitney U); open symbols denote electrodes that did not meet our significance threshold. This maze complexity effect also appears in the average power spectra for long (blue plots) vs. short (green plots) mazes, shown for selected electrodes (**d–g**).

- Bland, B. H. The physiology and pharmacology of hippocampal formation theta rhythms. *Prog. Neurobiol.* **26**, 1–54 (1986).
- Stewart, M. & Fox, S. E. Hippocampal theta activity in monkeys. *Brain Res.* **538**, 59–63 (1991).
- Huerta, P. T. & Lisman, J. E. Heightened synaptic plasticity of hippocampal CA1 neurons during a cholinergically induced rhythmic state. *Nature* **364**, 723–725 (1993).
- Buzsáki, G. Two-stage model of memory trace formation: role for “noisy” brain states. *Neuroscience* **31**, 551–570 (1989).
- Sohal, V. S. & Hasselmo, M. E. GABA(B) modulation improves sequence disambiguation in computational models of hippocampal region CA3. *Hippocampus* **8**, 171–193 (1998).
- Buzsáki, G. The hippocampo-neocortical dialogue. *Cereb. Cortex* **6**, 81–92 (1996).
- Lisman, J. E. & Idiart, M. A. Storage of 7+/-2 short-term memories in oscillatory subcycles. *Science* **267**, 1512–1515 (1995).
- Buzsáki, G. Memory consolidation during sleep: a neurophysiological perspective. *J. Sleep Res.* **7**, 17–23 (1998).
- Jensen, O. & Lisman, J. E. An oscillatory short-term memory buffer model can account for data on the Sternberg task. *J. Neurosci.* **18**, 10688–10699 (1998).
- Borst, J. G., Leung, L. W. & MacFabe, D. F. Electrical activity of the cingulate cortex. II. Cholinergic modulation. *Brain Res.* **407**, 81–93 (1987).
- Slawinska, U. & Kasicki, S. Theta-like rhythm in depth EEG activity of hypothalamic areas during spontaneous or electrically induced locomotion in the rat. *Brain Res.* **678**, 117–126 (1995).
- Routtenberg, A. & Taub, F. Hippocampus and superior colliculus: congruent EEG activity demonstrated by a simple measure. *Behav. Biol.* **8**, 801–805 (1973).
- Halgren, E., Babb, T. L. & Crandall, P. H. Human hippocampal formation EEG desynchronizes during attentiveness and movement. *Electroencephalogr. Clin. Neurophysiol.* **44**, 778–781 (1978).
- Arnolds, D. E. A. T., Lopes Da Silva, F. H., Aitink, J. W., Kamp, A. & Boeijinga, P. The spectral properties of hippocampal EEG related to behaviour in man. *Electroencephalogr. Clin. Neurophysiol.* **50**, 324–328 (1980).
- Tesche, C. D. Non-invasive detection of ongoing neuronal population activity in normal human hippocampus. *Brain Res.* **749**, 53–60 (1997).
- Aguirre, G. K., Detre, J. A., Alsop, D. C. & D’Esposito, M. The parahippocampus subserves topographical learning in man. *Cerebral Cortex* **6**, 823–829 (1996).
- Maguire, E. A. *et al.* Knowing where and getting there: a human navigation network. *Science* **280**, 921–924 (1998).
- Epstein, R. & Kanwisher, N. A cortical representation of the local visual environment. *Nature* **392**, 598–601 (1998).
- Maguire, E. A., Frackowiak, S. J. & Frith, C. D. Learning to find your way: a role for the human hippocampal formation. *Proc. R. Soc. Lond. B* **263**, 1745–1750 (1996).
- Maguire, E. A., Frackowiak, S. J. & Frith, C. J. Recalling routes around London: activation of the right hippocampus in taxi drivers. *J. Neurosci.* **17**, 7103–7110 (1997).
- Sperling, M. R. Clinical challenges in invasive monitoring in epilepsy surgery. *Epilepsia* (Suppl.) **38**, S6–S12 (1997).
- Talairach, J. & Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain* (Verlag, Stuttgart, 1988).
- Tulving, E. *Elements of Episodic Memory* (Oxford Univ. Press, New York, 1983).
- Grossmann, A. & Morlet, J. *Mathematics + Physics* Vol. 1 (World Scientific, Singapore, 1985).

Acknowledgements. This work was supported by NIH. We thank J. Lisman for encouraging our search for theta; J. Lisman and O. Jensen for helpful discussions during the course of this research; and E. Tulving, E. Marder, L. Abbott and E. Menschik for helpful comments on a previous version of this manuscript. We acknowledge the cooperation of colleagues in the Children’s Hospital Epilepsy program including P. M. Black, B. Bourgeois, J. Duffy and L. Kull. Finally, we thank the patients and their families for their participation and support.

Correspondence and requests for materials should be addressed to M.J.K. (e-mail: kahana@brandeis.edu).

The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein

Pei-Jung Lu*, Gerburg Wulf*, Xiao Zhen Zhou*, Peter Davies† & Kun Ping Lu*

* Cancer Biology Program, Division of Hematology/Oncology, Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts 02215, USA

† Department of Pathology, Albert Einstein College of Medicine, Bronx, New York 10461, USA

One of the neuropathological hallmarks of Alzheimer’s disease is the neurofibrillary tangle, which contains paired helical filaments (PHFs) composed of the microtubule-associated protein tau^{1,2}. Tau is hyperphosphorylated in PHFs^{3–5}, and phosphorylation of tau abolishes its ability to bind microtubules and promote microtubule assembly^{6,7}. Restoring the function of phosphorylated tau might prevent or reverse PHF formation in Alzheimer’s disease. Phosphorylation on a serine or threonine that precedes proline (pS/T–P) alters the rate of prolyl isomerization and creates a binding site for the WW domain of the prolyl isomerase Pin1 (refs 8–14). Pin1 specifically isomerizes pS/T–P bonds and

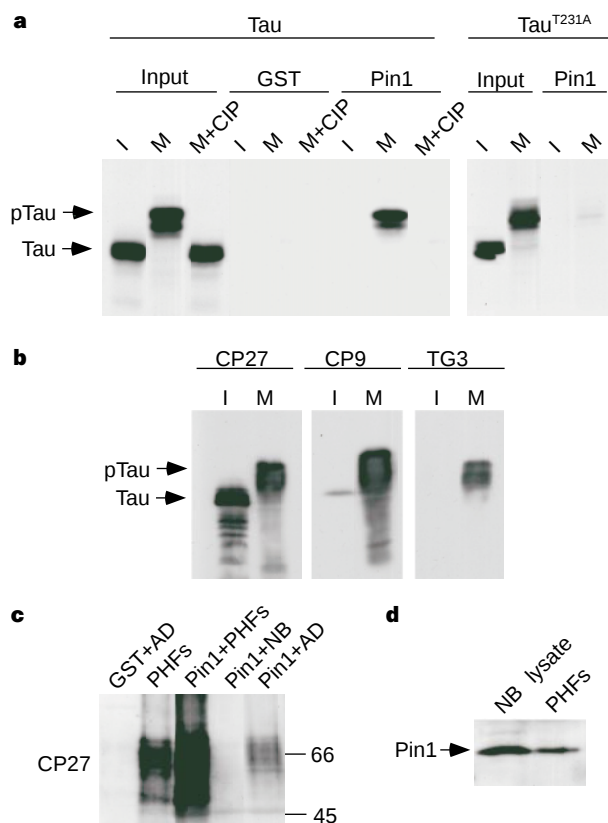


Figure 1 Interaction of Pin1 with pTau and Alzheimer’s disease tau, but not with pTau^{T231A}. **a**, *In vitro* translated tau and its mutant were incubated with interphase (I), mitotic extracts (M) or mitotic extracts followed by dephosphorylation with phosphatase (M + CIP), and then separated on SDS-PAGE directly (input) or after purification with GST (GST) or GST-Pin1 (Pin1) beads. **b**, Tau was phosphorylated by I or M extracts and immunoblotted with monoclonal antibodies with different specificity: CP27, all tau forms; CP9, pT231 tau; TG3, an Alzheimer-specific conformation of pT231 tau. **c**, Beads containing GST-Pin1 or GST were incubated with normal brain (NB) or Alzheimer’s disease (AD) brain extracts, or PHFs, and bead-associated proteins were immunoblotted with CP27. Lane 1, GST + AD; lane 2, PHFs; lane 3, GST-Pin1 + PHFs; lane 4, GST-Pin1 + NB; lane 5, GST-Pin1 + AD. **d**, PHFs were purified from AD brain extracts and immunoblotted with Pin1 antibodies. Pin1 present in normal brain extracts was used as a control.

regulates the function of mitotic phosphoproteins^{8–10,12}. Here we show that Pin1 binds to only one pT–P motif in tau and co-purifies with PHFs, resulting in depletion of soluble Pin1 in the brains of Alzheimer’s disease patients. Pin1 can restore the ability of phosphorylated tau to bind microtubules and promote microtubule assembly *in vitro*. As depletion of Pin1 induces mitotic arrest and apoptotic cell death⁸, sequestration of Pin1 into PHFs may contribute to neuronal death. These findings provide a new insight into the pathogenesis of Alzheimer’s disease.

Pin1 is a highly conserved and essential mitotic regulator^{10,12–15}. It contains an amino-terminal WW domain, a phosphoserine-binding module interacting with specific pS/T–P motifs¹³, and a unique carboxy-terminal prolyl isomerase domain that specifically isomerizes pS/T–P bonds¹⁰. Pin1 binds and regulates the function of a subset of mitotic phosphoproteins, most of which are also recognized by MPM-2, a mitosis- and phosphorylation-specific monoclonal antibody^{10,12–15}. As tau is an MPM-2 antigen that is phosphorylated on multiple S/T–P motifs during mitosis¹⁶, we tested whether Pin1 binds tau. Pin1 bound tau only after mitosis-specific phosphorylation, and the binding was abolished by dephosphorylation (Fig. 1a). Thus, Pin1 binds tau in a mitosis-specific and phosphorylation-dependent manner, as shown for other binding proteins¹². Mitotic events are aberrantly activated in the brains

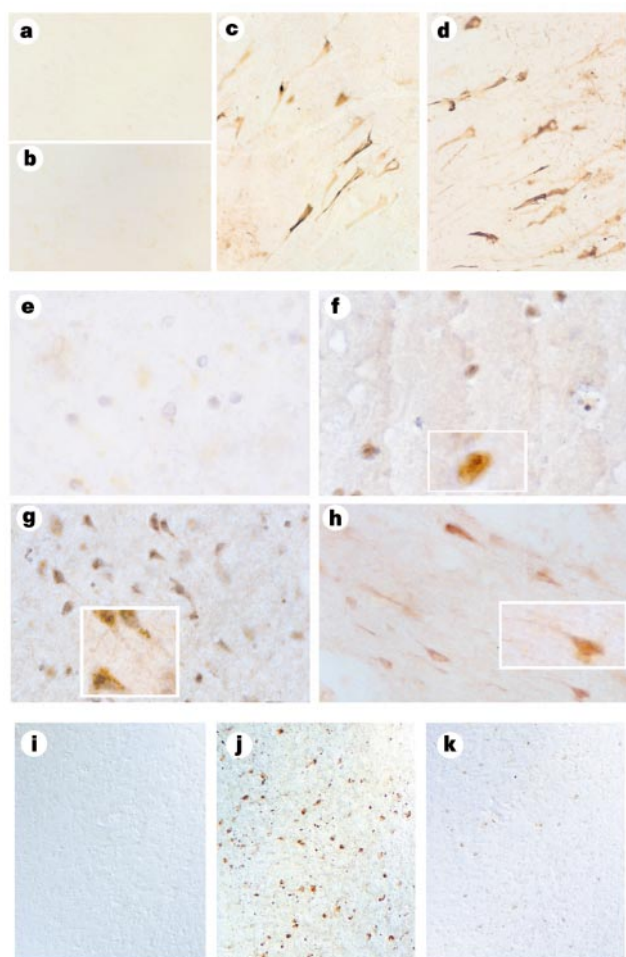


Figure 2 Localization of Pin1 in normal and AD brains and competition of Pin1 binding by pT231 antibody. **a-d**, Rehydrated fixed sections were incubated with 0.5 μ M Pin1 or buffer control, followed by immunostaining with Pin1 antibodies or TG3. **a**, NB + Pin1; **b**, AD + buffer; **c**, AD + Pin1; **d**, AD section stained with TG3. **e-h**, Sections from normal and AD brains were first subjected to an antigen-retrieval procedure, followed by immunostaining with Pin1 antibodies or TG3. **e**, Pin1-specific antibodies were depleted before immunostaining; **f**, normal brains stained with Pin1 antibodies; **g**, AD brain stained with Pin1 antibodies; **h**, AD brain stained with TG3. Sections were counter-stained with haematoxylin and eosin. Inserts show higher magnification pictures. **i-k**, Rehydrated fixed sections were incubated with buffer control (**i**), 0.5 μ M Pin1 (**j**) or first CP9 antibody and then Pin1 (**k**), followed by immunostaining with Pin1 antibodies.

of patients with Alzheimer's disease (AD), including re-expression of Cdc2 kinase and cyclin B¹⁷⁻¹⁹. The phosphorylation patterns of tau in mitotic cells and AD brains are strikingly similar^{17,18,20,21}. We also confirmed that mitotically phosphorylated tau (pTau) was recognized by AD-specific, phosphorylation-dependent tau monoclonal antibodies, including CP9, TG3 and PHF1 (Fig. 1b)^{17,18,20,21}. These results indicate that the common S/T-P motifs of tau are phosphorylated in normal mitotic cells and in AD brains. We therefore considered that Pin1 might bind and regulate the function of tau in the AD brain.

We first investigated whether Pin1 binds tau in AD brain extracts by using a glutathione S-transferase (GST)-Pin1 pulldown assay¹². Glutathione beads containing GST or GST-Pin1 were incubated with normal or AD brain extracts, or purified PHFs²². Pin1-binding proteins were then analysed by immunoblotting with the antibody CP27, which recognizes all isoforms of tau. Figure 1c shows that GST-Pin1, but not GST, bound tau in AD extracts or PHFs. In contrast, Pin1 did not bind any tau in extracts from age-matched

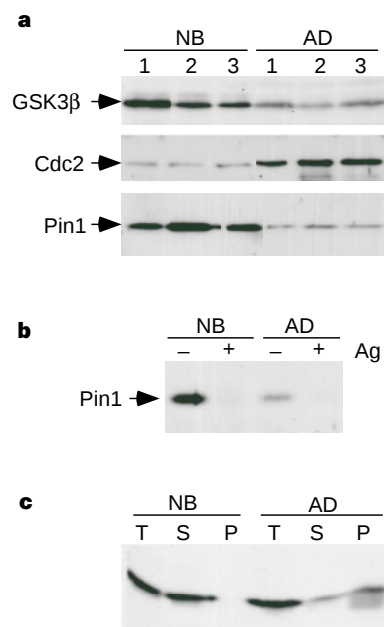


Figure 3 Depletion of soluble Pin1 in AD brains. **a**, Brain tissues were homogenized and soluble protein was analysed by immunoblotting with antibodies against GSK3 β , Cdc2 and Pin1. Lanes 1-3, three representative normal individuals; lanes 4-6, three representative Alzheimer's disease patients. **b**, Soluble proteins isolated from normal or AD brain extracts were immunoprecipitated using Pin1 antibodies in the absence (-) or presence (+) of excess Pin1 antigen, followed by immunoblotting analysis with Pin1 antibodies. **c**, Homogenized lysates (T) of normal and AD brain tissues were separated into soluble (S) and insoluble (P) fractions by centrifugation, and immunoblotted with Pin1 antibodies.

normal brains (Fig. 1c). To examine whether Pin1 forms a complex with PHFs *in vivo*, PHFs were purified from AD brains using a series of procedures, including immunoaffinity chromatography²², and then analysed by immunoblotting with anti-Pin1 antibodies. Pin1 was detected in PHFs purified from all six AD brains examined (Fig. 1d, data not shown). These results indicate that Pin1 binds PHFs in Alzheimer's disease brains.

To confirm that Pin1 has specific affinity only for PHFs, we incubated brain sections with recombinant Pin1 and then stained the sections with affinity-purified Pin1 antibodies to localize the bound Pin1. If no Pin1 was added, no immunoreactive signal was observed (Fig. 2b), indicating that Pin1 antibodies do not recognize endogenous Pin1 under the conditions used. However, if normal and AD brain sections were incubated with Pin1 protein, robust Pin1 binding was detected in the cytoplasm of neurons in AD brain sections (Fig. 2c) but not in normal brain (Fig. 2a). Pin1 specifically bound to neurofibrillary tangles and neurites (Fig. 2c), which were also detected by staining with TG3 (Fig. 2d), a monoclonal antibody recognizing the AD-specific conformation of tau phosphorylated on Thr 231 (ref. 23). Thus, exogenous Pin1 specifically binds the neurofibrillary tangles in neurons.

Given that Pin1 has a high affinity for the tangles and co-purifies with PHFs, it is critical to show that Pin1 binds PHFs *in vivo*. We first subjected the brain sections to an antigen-retrieval procedure and saw strong immunoreactivity with the Pin1 antibodies both in normal and AD brain sections (Fig. 2f, g). To ensure that these signals represented Pin1, the Pin1-specific antibodies were first depleted using GST-Pin1 beads and then used for immunostaining. The Pin1-depleted antibodies showed no specific immunoreactivity (Fig. 3e). Very different patterns of Pin1 localization were seen in normal and AD brain sections, Pin1 was mainly localized in the nuclei of neurons in normal brain and was also detected in neuronal

nuclei in AD brains (Fig. 2f, g), consistent with reports that Pin1 is primarily localized in the nucleus of HeLa cells⁸. However, in AD brains, robust Pin1 immunostaining was also found in the cytoplasm of neurons, specifically at the tangle structure that was also recognized by TG3 (ref. 23) (Fig. 2g, h). Therefore, both exogenous and endogenous Pin1 specifically localize to the neurofibrillary tangles in Alzheimer's disease brains.

Binding of Pin1 to PHFs might trap Pin1 in the tangles, leading to depletion of soluble Pin1 in neurons. To test this possibility, we compared the amounts of Pin1 and two tau kinases, GSK3 β and Cdc2, in the soluble fraction of AD and normal brain extracts. When compared with age-matched normal brains, GSK3 β was only slightly reduced ($40 \pm 11\%$, $n = 6$) (Fig. 3a, top), but Cdc2 levels were significantly increased by ~ 5 fold in diseased brains ($547 \pm 87\%$, $n = 6$) (Fig. 3a, middle). These findings are consistent with previous findings that Cdc2, but not GSK3 β , is abnormally elevated in the brain in Alzheimer's disease¹⁸. Levels of soluble Pin1 in all diseased brains examined were much lower than those in normal brains (Fig. 3a, bottom), being reduced by an average of ~ 5 fold ($22.4 \pm 3.4\%$, $n = 6$). This Pin1 decrease was further confirmed by immunoprecipitation analysis (Fig. 3b). If Pin1 is trapped in the tangles, most of the Pin1 in diseased brains would be in the insoluble fraction. As shown in Fig. 3c, although the total levels of Pin1 were not significantly different, most Pin1 was detected in the soluble fraction in normal brain, but in the insoluble fraction in AD brain (Fig. 3c). Together with the results on Pin1 localization, these data indicate that Pin1 is relocated to the neurofibrillary tangles in Alzheimer's disease, resulting in depletion of soluble Pin1.

Pin1 binds mitotic phosphoproteins through its N-terminal WW domain¹³. To identify the Pin1 binding site on tau, we synthesized a series of phosphorylated and non-phosphorylated peptides that covered most known tau phosphorylation sites, and then tested their ability to bind Pin1 using the ELISA (enzyme-linked immunosorbent) assay²³ (Table 1). Although little or no binding was detected between Pin1 and most phosphopeptides, Pin1 exhibited specific and high-affinity binding to one tau peptide containing pT231 (Table 1), with a dissociation constant of ~ 40 nM (Fig. 4a). In contrast, no binding was observed between Pin1 and the non-phosphorylated tau (Fig. 4a), demonstrating that T231 phosphorylation is required for Pin1 binding of tau. To determine whether the WW domain of Pin1 is responsible for binding, the mutant Pin1^{Y23A} was used, which contains a single alanine substitution at the critical Tyr 23 in the WW domain, resulting in a loss of the phosphoserine-binding activity¹³. Pin1^{Y23A} showed much less binding to pT231 peptide (Table 1). The residual binding might be due to binding of the pT231 peptide to the much lower affinity isomerase domain of Pin1¹³. These results indicate that the WW domain mediates Pin1 binding to the pT231 sequence of tau.

Table 1 Identification of the Pin1-binding site in tau

Pin1	Tau peptides	Binding (OD at 405 nm)
Pin1	DAGLKEpSPLQTPTE (pS46)	0.00
	TRIPAKpTPAPKT (pT175)	0.00
	GYSSPGpSPGITPGSR (pS202)	0.08
	SRSRTpSLPTPTPT (pS214)	0.00
	KVSVVRTPPKSPS (T231)	0.00
	KVSVVRTPPKSPS (pT231)	1.46
	VRTPPKpSPSSAKSR (pS235)	0.11
	VQSKIGpSLDNITH (pS356)	0.00
	GSLDNIPpTHVPGGG (pT361)	0.00
	TSPPIHLpSNVSSTG (pS409)	0.00
	PRHLNVpSSTGSIDMV (pS412)	0.02
	PRHLNVpSSTGSIDMV (pS413)	0.00
	NVSSTGpSIDMVDS (pS416)	0.00
	SIDMVDPpSQLATL (pS422)	0.00
Pin1 ^{Y23A}	KVAVVRpTPPKSPS (pT231)	0.18

A series of tau peptides were synthesized and labelled with an N-terminal biotin tag, then immobilized on 96-well plates. Pin1 and its mutant were added to the plates and extensively washed, followed by detection of Pin1 binding by ELISA analysis using affinity-purified Pin1 antibodies.

Phosphorylation of tau on Thr 231 (pT231 tau) has been well documented in Alzheimer's disease brains and can be recognized by antibodies including CP9 (refs 16, 18, 21, 23). To test whether Pin1 interacts with pT231-tau, we used GST-Pin1 beads to isolate tau from Alzheimer's disease brain extracts or PHFs, and then detected Thr 231 phosphorylation using CP9. Tau isolated by Pin1 beads was strongly immunoreactive with CP9 (Fig. 4b), indicating that Pin1 binds pT231 tau and that this binding does not induce dephosphorylation of pT231. As T231 is readily phosphorylated by Cdc2 kinase *in vitro*^{17,18,23}, we examined whether Pin1 binds tau phosphorylated by Cdc2. Pin1 and its WW domain, but not its isomerase domain, bound Cdc2-pTau (Fig. 4c). These results support the idea that Pin1 binds pT231 tau through its WW domain.

To show that pT231 is both necessary and sufficient for mediating the interaction between Pin1 and tau, we tested whether three different pT231-specific antibodies, CP9, CP17 and TG3, compete with Pin1 for binding to PHFs in brain sections. Each of the antibodies strongly competed with Pin1 for binding to PHFs (Fig. 2i–k and data not shown), showing that pT231 is needed for Pin1 to bind PHFs. We next replaced Thr 231 with the non-phosphorylatable alanine

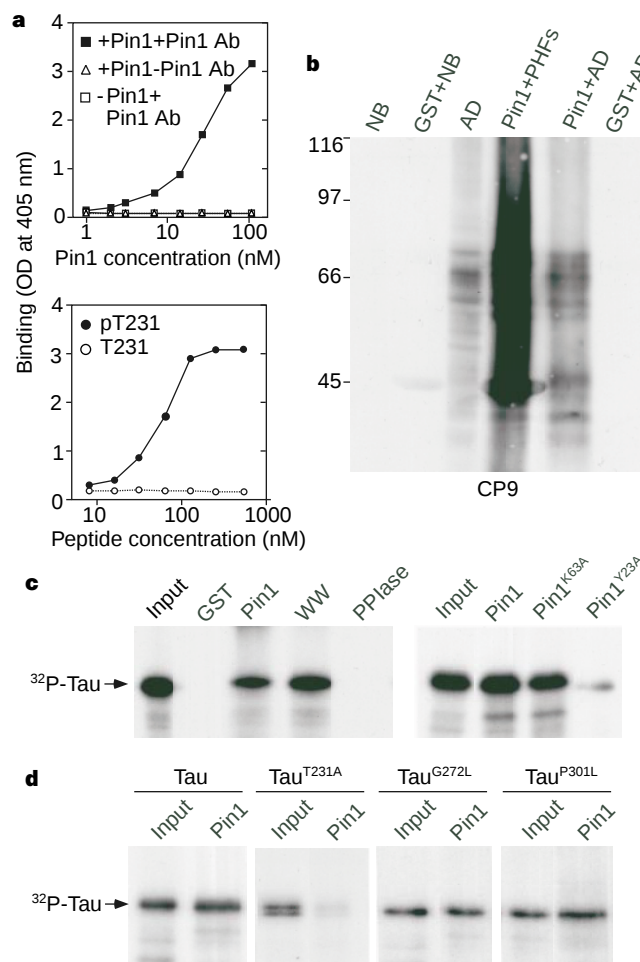


Figure 4 Characterization of the Pin1-tau interaction. **a**, Different concentrations of Pin1 were incubated with a fixed concentration of T231 or pT231 tau peptide (top) or vice versa (bottom), followed by ELISA assay with Pin1 antibodies. **b**, Beads containing GST or GST-Pin1 were incubated with PHFs, or normal or Alzheimer's disease brain extracts, followed by immunoblotting analysis with CP9. **c**, Tau was phosphorylated by Cdc2 kinase in the presence of [γ -³²P]ATP, and then subjected to GST pull-down assay with Pin1 or its mutants. WW and PPlase are the N-terminal WW domain and C-terminal catalytic domain, respectively. **d**, Tau and its mutants were phosphorylated by cyclin B/Cdc2, and then either separated on SDS-PAGE directly (input) or after affinity purification using GST-Pin1 beads (Pin1).

(tau^{T231A}) to rule out the possibility that Pin1 binds tau through other pS/T sequences. When incubated with either mitotic extracts or purified Cdc2 kinase, tau^{T231A} was still phosphorylated (Figs 1a, 4d). However, Pin1 almost completely failed to bind tau^{T231A} that was phosphorylated by either method (Figs 1a, 4d). In contrast, two other tau mutations, G272L and P301L, which are found in patients with hereditary forms of fronto-temporal dementia (FTDP-17)^{24–26}, did not affect Pin1 binding of pTau (Fig. 4d). Thus, pT231 is the only Pin1-binding site in tau. In addition, phosphorylation of tau^{T231A} by Cdc2 was significantly reduced. This result confirms previous findings that T231 is an important Cdc2 phosphorylation site in tau, and is also consistent with Pin1 binding mitotically pTau (Fig. 1a) and being sequestered on to PHFs in AD brains, where Cdc2 is upregulated (Fig. 3).

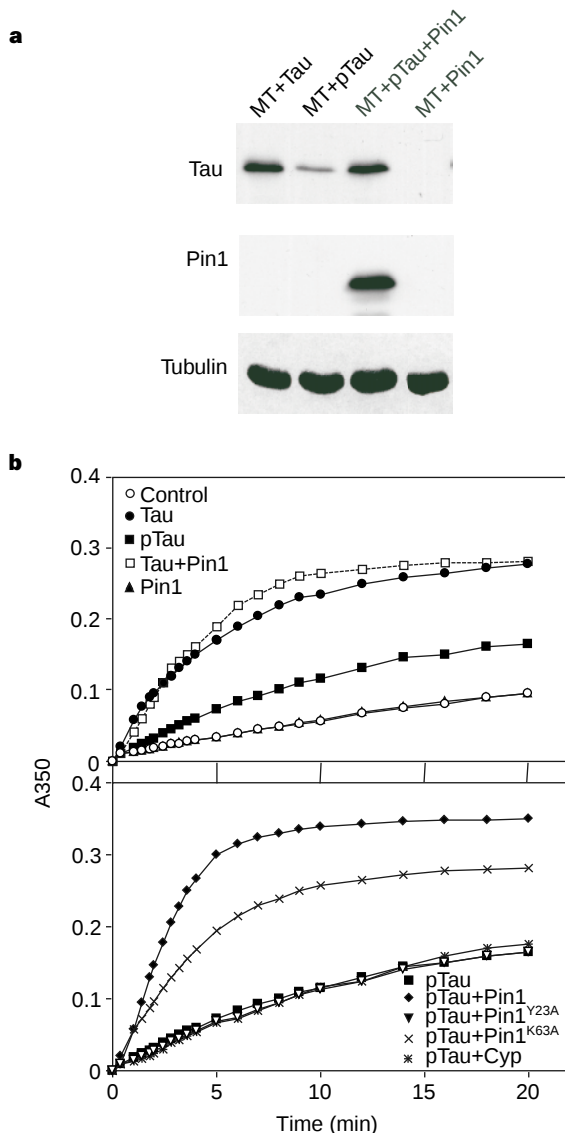


Figure 5 Restoration of function of pTau by Pin1, but not its mutants or cyclophilin. **a**, Taxol-stabilized microtubules were incubated with non-phosphorylated or Cdc2-pTau in the presence or absence of Pin1 or directly incubated with Pin1. A sucrose cushion was underlain and the bound tau was separated by centrifugation, followed by immunoblotting analysis with the tau antibody CP27 or Pin1 antibodies. The membranes were also stained with Coomassie blue, as shown in the lower panel. **b**, Top: rate of microtubule assembly promoted by tau in the presence and absence of Pin1, or by Cdc2-pTau. Control, tubulin only; Tau, tubulin + tau; pTau, tubulin + tau phosphorylated by Cdc2; Tau + Pin1, tubulin + tau + Pin1; Pin1, tubulin + Pin1. Bottom: rate of microtubule assembly promoted by Cdc2-pTau with or without of Pin1, its mutants or cyclophilin (Cyp), as indicated.

The high-affinity interaction between Pin1 and pTau indicates that Pin1 might affect the biological activity of pTau. When it is phosphorylated by many protein kinases, including Cdc2, tau loses its ability to bind microtubules and promote their assembly^{6,7,27,28}. To examine whether Pin1 affects the ability of pTau to bind microtubules, we generated pTau *in vitro* using purified Cdc2 (refs 17, 18), and determined its ability to bind Taxol-stabilized microtubules with or without Pin1. Although Cdc2 phosphorylation disrupted the ability of tau to bind microtubules; the binding was fully restored by preincubation with Pin1 (Fig. 5a). Furthermore, Pin1 was detected in the fraction of tau-bound microtubules (Fig. 5a). However, no Pin1 was detected in the microtubule fraction if pTau was not added (Fig. 5a), indicating that Pin1 does not bind microtubules directly. Thus, Pin1 binds pTau and restores its ability to bind microtubules.

We next assessed the effect of Pin1 on the ability of pTau to promote microtubule assembly by using light-scattering assays^{27,28}. The rate of the turbidity change was minimal without tau, but was markedly increased if tau was added (Fig. 5b, top). This increase was almost abolished if tau was phosphorylated by Cdc2 (Fig. 5b), confirming that phosphorylation of tau by Cdc2 disrupts its ability to promote microtubule assembly^{6,7,28}. Importantly, Pin1 fully restored the ability of Cdc2-phosphorylated tau to promote microtubule assembly in a dose-dependent manner, although Pin1 had no effect on non-phosphorylated tau (Fig. 5b, bottom, data not shown). To evaluate the specificity of the effect of Pin1, we first used the isomerase cyclophilin and the pT231-specific antibody CP9 in the assay. Neither had any detectable effects on pTau (Fig. 5b, bottom, and data not shown). Next, we introduced point mutations into Pin1 specifically to affect its phosphoprotein-binding or isomerase activity. The substitution of Tyr 23 with alanine in the WW domain does not affect the isomerase activity, but disrupts the ability of Pin1 to bind phosphoproteins¹³ including pTau (Fig. 4c). This Y23A mutation also abolished the ability of Pin1 to restore the function of pTau (Fig. 5b, bottom), indicating an essential role for Pin1 binding. Conversely, introducing alanine into Lys 63, an active-site residue conserved in all Pin1-related isomerases¹⁴, did not affect binding to pTau (Fig. 4c), but did reduce isomerase activity to about 10% of the wild-type level, as assayed with an *in vitro* peptide substrate¹⁰. This mutation also significantly reduced the ability of Pin1 to restore promotion of microtubule assembly by pTau (Fig. 5b, bottom), indicating a role for the isomerase activity in regulating tau function. These data show that Pin1 not only binds pTau, but also restores its biological activity.

Tau stabilizes the internal microtubular structure of neurons^{1,2}. The importance of tau is demonstrated by findings that tau mutations cause the inherited dementia FTDP-17 (refs 24–26). Tau is hyperphosphorylated in FTDP-17, as in Alzheimer's disease²⁹. Furthermore, some FTDP-17 mutations also disrupt the ability of tau to bind microtubules and promote their assembly³⁰, demonstrating a critical role for the tau-microtubule interaction. Pin1 binds Alzheimer's disease-associated tau and restores the ability of pTau to bind microtubules and promote their assembly. Pin1 binds pS/T–P motifs and isomerizes pS/T–P peptide bonds^{10,12,13}. The initial mutational analyses have revealed that each of these activities is critical for regulating tau function. This is consistent with the fact that both the WW domain and the isomerase domain are essential for Pin1 function *in vivo*^{8,13}. Although further work is needed to elucidate the exact role of the isomerase activity, it is possible that Pin1 binds and somehow alters the conformation of pS/T–P motifs in tau to regulate its biological function.

As depletion of Pin1 induces mitotic arrest and apoptosis⁸, Pin1 may be required to prevent abnormal activation of mitotic events in neurons and to control the function of phosphoproteins, such as tau. Abnormal activation of mitotic events would result in hyperphosphorylation of tau, which binds and sequesters Pin1, as seen in Alzheimer's disease brains^{17–19}. This may have two consequences.

First, hyperphosphorylation of tau may create more binding sites than the capacity of the available Pin1, as indicated by extra binding sites in PHFs for exogenous Pin1. This hyperphosphorylated tau cannot bind microtubules, and subsequently forms PHFs, affecting neuronal function^{1,2}. On the other hand, sequestration of Pin1 in PHFs depletes soluble Pin1, which itself might also have a deleterious effect on neurons that already contain mitotic phosphoproteins. Therefore, both depletion of Pin1 and formation of PHFs might contribute to neuronal loss in Alzheimer's disease. As the aggregates of hyperphosphorylated tau are also a common neuropathological feature of several other neuronal degenerative diseases, such as FTDP-17 (ref. 29), Pin1 may also be involved in these diseases. Future studies on the role of Pin1 in these neurodegenerative diseases, including the identification of possible Pin1 mutations, would help us to understand their aetiology. In addition, the prolyl isomerase Pin1 might be used to derive new therapies for Alzheimer's and related neurodegenerative diseases. □

Methods

Production and phosphorylation of tau. The longest tau isoform and its mutants, generated by PCR mutagenesis^{10,12}, were either synthesized by *in vitro* transcription and translation in the presence of ³⁵S-Met, or produced in bacteria as N-terminal His-tagged proteins, followed by purification on NTA-Ni columns, as described^{10,12}. To generate the interphase- and mitosis-specific phosphorylated form, we incubated tau proteins with *Xenopus* interphase and mitotic extracts, respectively¹². To prepare Cdc2 pTau, we incubated purified recombinant tau proteins with purified cyclin B/Cdc2 (UBI) for 6–12 h at room temperature in a buffer containing 500 μ M cold ATP, plus trace [³²P]ATP in some experiments^{13,18}.

Determination of Pin1-tau interaction and Pin1 levels. Recombinant and mutant Pin1 proteins were produced as N-terminal GST or His-tagged fusion proteins, as described^{12,13}. Pin1 binding to tau peptides was assayed using ELISA, as described²³, which gave lower K_d values than those obtained using a fluorescence-polarization binding assay¹³. The interaction between Pin1 and tau or its mutants was determined using a GST–Pin1 pulldown assay, as described¹². PHFs were purified by a series of procedures, including immunoaffinity chromatography with an anti-tau monoclonal antibody²², and Pin1 and tau antibodies (CP9, CP17, CP27, TG3 and PHF1) have been described^{12,23}. To determine levels of Pin1, brain tissues were sliced, cut into fine pieces and homogenized in buffer A (50 mM HEPES, pH 7.4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 5 mM MgCl₂, 1 mM EGTA, 1 mM DTT, 100 mM NaF, 2 mM Na₃VO₄ and various protease inhibitors). The homogenates were centrifuged at 100,000g at 4°C for 30 min and the supernatants were directly used for immunoprecipitations or immunoblotting analysis with Pin1 antibodies¹², or stored in aliquots at –80°C before assays. The pellets were boiled in SDS-sample buffer repeatedly before subjecting to immunoblotting analysis. All Alzheimer's disease cases used were clinically diagnosed with senile dementia of the Alzheimer's type and confirmed histopathologically to have plaques and tangles, whereas control brains were from age-matched normal individuals^{17,18,23}.

Immunocytochemistry. To detect the localization of exogenously added Pin1, 50 μ m sections were cut from formalin-fixed frontal cortex or hippocampus of human brains, and endogenous peroxidase activity was blocked with H₂O₂, followed by incubation with Pin1 at 0.5 μ M. For antibody competition experiments, monoclonal antibodies were added onto brain sections 1 h before adding Pin1. After extensive washing, sections were incubated with anti-Pin1 antibodies that had been purified using GST–Pin1 glutathione beads, and visualized by the immunoperoxidase staining protocol, as described²³. To localize endogenous Pin1, the fixed brain sections were first microwaved in an antigen-retrieval buffer (Biogenex), as described by the manufacturer, before undergoing immunostaining procedures.

Microtubule binding and assembly. The ability of tau to bind microtubules was determined as described^{27,28}. Briefly, microtubules were assembled from purified bovine tubulin (cytoskeleton) and stabilized by Taxol. The non-phosphorylated or Cdc2-phosphorylated recombinant tau (0.1 mg ml^{–1}) was incubated with Pin1 (0.1 mg ml^{–1}) at 35°C for 5 min before adding to the microtubules. Bound tau was isolated by centrifugation (50,000g) at 25°C for 20 min, followed by immunoblotting analysis using CP27 and Pin1 antibodies.

The ability of tau to promote microtubule assembly was determined using light-scattering assays^{27,28}, with some modifications. Briefly, microtubule assembly was initiated by incubating tubulin (2 mg ml^{–1}) with or without tau (0.05 mg ml^{–1}) in 80 mM PIPES, pH 6.8, 1 mM EGTA, 1 mM MgCl₂, 1 mM GTP, 20% glycerol at 35°C for 2 min. The mixture was then transferred to a 100 μ l cuvette and the rate of microtubule assembly was monitored at room temperature by following the turbidity increase at 350 nm. To examine the effect of added proteins, tau or Cdc2-pTau (0.05 mg ml^{–1}) was preincubated with Pin1, its mutants, CP9 or cyclophilin (Sigma) (0.05 mg ml^{–1}) at 35°C for 5 min before the microtubule assembly assays. To determine stoichiometric amounts, Pin1 concentrations were reduced from 0.05 mg ml^{–1} to 0.01 mg ml^{–1}, and similar results were obtained. Each experiment was repeated at least three times, with similar results. Results using GST–Pin1 or His–Pin1 were similar, indicating that the N-terminal tags have no effect on the microtubule assembly assay.

Received 7 April; accepted 4 May 1999.

- Kosik, K. S., Qiu, W. Q. & Greenberg, S. Cellular signaling pathways and cytoskeletal organization. *Ann. NY Acad. Sci.* **777**, 114–120 (1996).
- Spillantini, M. G. & Goedert, M. Tau protein pathology in neurodegenerative diseases. *Trends Neurosci.* **21**, 428–433 (1998).
- Lee, V. M., Balin, B. J., Otvos, L. Jr & Trojanowski, J. Q. A68: a major subunit of paired helical filaments and derivatized forms of normal Tau. *Science* **251**, 675–678 (1991).
- Goedert, M., Spillantini, M. G., Cairns, N. J. & Crowther, R. A. Tau proteins of Alzheimer paired helical filaments: abnormal phosphorylation of all six brain isoforms. *Neuron* **8**, 159–168 (1992).
- Greenberg, S. G., Davies, P., Schein, J. D. & Binder, L. I. Hydrofluoric acid-treated tau PHF proteins display the same biochemical properties as normal tau. *J. Biol. Chem.* **267**, 564–569 (1992).
- Bramblett, G. T. et al. Abnormal tau phosphorylation at Ser396 in Alzheimer's disease recapitulates development and contributes to reduced microtubule binding. *Neuron* **10**, 1089–1099 (1993).
- Yoshida, H. & Ihara, Y. Tau in paired helical filaments is functionally distinct from fetal tau: assembly incompetence of paired helical filament-tau. *J. Neurochem.* **61**, 1183–1186 (1993).
- Lu, K. P., Hanes, S. D. & Hunter, T. A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **380**, 544–547 (1996).
- Ranganathan, R., Lu, K. P., Hunter, T. & Noel, J. P. Structural and functional analysis of the mitotic peptidyl-prolyl isomerase Pin1 suggests that substrate recognition is phosphorylation dependent. *Cell* **89**, 875–886 (1997).
- Yaffe, M. B. et al. Sequence-specific and phosphorylation-dependent proline isomerization: A potential mitotic regulatory mechanism. *Science* **278**, 1957–1960 (1997).
- Schutkowski, M. et al. Role of phosphorylation in determining the backbone dynamics of the serine/threonine-proline motif and Pin1 substrate recognition. *Biochemistry* **37**, 5566–5575 (1998).
- Shen, M., Stukenberg, P. T., Kirschner, M. W. & Lu, K. P. The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins. *Genes Dev.* **12**, 706–720 (1998).
- Lu, P. J., Zhou, X. Z., Shen, M. & Lu, K. P. Function of WW domains as phosphoserine or phosphothreonine-binding modules. *Science* **283**, 1325–1328 (1999).
- Lu, K. P. Phosphorylation-dependent prolyl isomerization: a novel cell cycle regulatory mechanism. *Prog. Cell Cycle Res.* (in press).
- Crenshaw, D. G., Yang, J., Means, A. R. & Kornbluth, S. The mitotic peptidyl-prolyl isomerase, Pin1, interacts with Cdc25 and Plx1. *EMBO J.* **17**, 1315–1327 (1998).
- Illenberger, S. et al. The endogenous and cell cycle-dependent phosphorylation of tau protein in living cells: implications for Alzheimer's disease. *Mol. Biol. Cell* **9**, 1495–1512 (1998).
- Vincent, I., Rosado, M. & Davies, P. Mitotic mechanisms in Alzheimer's disease? *J. Cell Biol.* **132**, 413–425 (1996).
- Vincent, I., Jicha, G., Rosado, M. & Dickson, D. W. Aberrant expression of mitotic cdc2/cyclin B1 kinase in degenerating neurons of Alzheimer's disease brain. *J. Neurosci.* **17**, 3588–3598 (1997).
- Nagy, Z., Esiri, M. M. & Smith, A. D. Expression of cell division markers in the hippocampus in Alzheimer's disease and other neurodegenerative conditions. *Acta Neuropathol. (Berl.)* **93**, 294–300 (1997).
- Kondratieff, C. M. & Vandre, D. D. Alzheimer's disease neurofibrillary tangles contain mitosis-specific phosphopeptides. *J. Neurochem.* **67**, 2405–2416 (1996).
- Preuss, U. & Mandelkow, E. M. Mitotic phosphorylation of tau protein in neuronal cell lines resembles phosphorylation in Alzheimer's disease. *Eur. J. Cell Biol.* **76**, 176–184 (1998).
- Vincent, I. J. & Davies, P. A protein kinase associated with paired helical filaments in Alzheimer disease. *Proc. Natl Acad. Sci. USA* **89**, 2878–2882 (1992).
- Jicha, G. A. et al. A conformation- and phosphorylation-dependent antibody recognizing the paired helical filaments of Alzheimer's disease. *J. Neurochem.* **69**, 2087–2095 (1997).
- Hutton, M. et al. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **393**, 702–705 (1998).
- Clark, L. N. et al. Pathogenic implications of mutations in the tau gene in palidoponto-nigral degeneration and related neurodegenerative disorders linked to chromosome 17. *Proc. Natl Acad. Sci. USA* **95**, 13103–13107 (1998).
- Spillantini, M. G. et al. Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc. Natl Acad. Sci. USA* **95**, 7737–7741 (1998).
- Iqbal, K., Zaidi, T., Bancher, C. & Grundke-Iqbal, I. Alzheimer paired helical filaments. Restoration of the biological activity by dephosphorylation. *FEBS Lett.* **349**, 104–108 (1994).
- Alonso, A. C., Zaidi, T., Grundke-Iqbal, I. & Iqbal, K. Role of abnormally phosphorylated tau in the breakdown of microtubules in Alzheimer disease. *Proc. Natl Acad. Sci. USA* **91**, 5562–5566 (1994).
- Spillantini, M. G., Bird, T. D. & Ghetti, B. Frontotemporal dementia and Parkinsonism linked to chromosome 17: a new group of tauopathies. *Brain Pathol.* **8**, 387–402 (1998).
- Hong, M. et al. Mutation-specific functional impairments in distinct tau isoforms of hereditary FTDP-17. *Science* **282**, 1914–1917 (1998).

Acknowledgements. We thank B. Neel, L. Cantley, C. S. Chen, G. Lee and X. D. Fu for constructive comments. P.-J.L. and G.W. are fellows of Leukemia Society of America and DOE Breast Cancer Research Program, respectively. K.P.L. is a Pew Scholar and a Leukemia Society of America Scholar. This study was supported by NIH grants to P.D. and K.P.L.

Correspondence and requests for materials should be addressed to K.P.L. (e-mail: klu@caregroup.harvard.edu).

Characterization of the human cysteinyl leukotriene CysLT₁ receptor

Kevin R. Lynch*, Gary P. O'Neill†, Qingyun Liu§, Dong-Soon Im*, Nicole Sawyer†, Kathleen M. Metters†, Nathalie Coulombe†, Mark Abramovitz†, David J. Figueroa§, Zhizhen Zeng§, Brett M. Connolly§, Chang Bai§, Christopher P. Austin§, Anne Chateaufort†, Rino Stocco†, Gillian M. Greig†, Stacia Kargman†, Shelley B. Hooks*, Elizabeth Hosfield*, David L. Williams Jr†, Anthony W. Ford-Hutchinson‡, C. Thomas Caskey§ & Jilly F. Evans§

* Department of Pharmacology, University of Virginia Health Sciences Center, Charlottesville, Virginia 22908, USA

† Department of Biochemistry and Molecular Biology, Merck Frosst Canada & Co., PO Box 1005, Pointe Claire-Dorval, Quebec, Canada H9R 4P8

‡ Department of Pharmacology and § Department of Human Genetics, Merck & Co., Summerville Pike, PO Box 4, West Point, Pennsylvania 19486, USA

The cysteinyl leukotrienes—leukotriene C₄(LTC₄), leukotriene D₄(LTD₄) and leukotriene E₄(LTE₄)—are important mediators of human bronchial asthma^{1–3}. Pharmacological studies have determined that cysteinyl leukotrienes activate at least two receptors, designated CysLT₁ and CysLT₂ (refs 4–6). The CysLT₁-selective antagonists, such as montelukast (Singulair)^{7–10}, zafirlukast (Accolate)¹¹ and pranlukast (Onon)¹², are important in the treatment of asthma. Previous biochemical characterization of CysLT₁ antagonists and the CysLT₁ receptor has been in membrane preparations from tissues enriched for this receptor¹³. Here we report the molecular and pharmacological characterization of the cloned human CysLT₁ receptor. We describe the functional activation (calcium mobilization) of this receptor by LTD₄ and LTC₄,

and competition for radiolabelled LTD₄ binding to this receptor by the cysteinyl leukotrienes and three structurally distinct classes of CysLT₁-receptor antagonists. We detected CysLT₁-receptor messenger RNA in spleen, peripheral blood leukocytes and lung. In normal human lung, expression of the CysLT₁-receptor mRNA was confined to smooth muscle cells and tissue macrophages. Finally, we mapped the human CysLT₁-receptor gene to the X chromosome.

CysLT₁-receptor activation by LTD₄ results in contraction and proliferation of smooth muscle, oedema, eosinophil migration and damage to the mucus layer in the lung^{1–3}. An elegant biochemical characterization of the photoaffinity-labelled CysLT₁ receptor from guinea pig lung shows that it is a glycosylated G-protein-coupled receptor (GPCR)¹³. Intense efforts to clone this receptor, either by homology screening or expression cloning, have so far been unsuccessful. The only reported leukotriene-receptor sequence is for the BLT receptor that is activated by the proinflammatory mediator leukotriene B₄ (LTB₄)^{14,15}.

As part of a programme to identify ligands for orphan GPCRs we identified an expressed sequence tag (EST) entry (GenBank accession number AA832466) that had 32% amino-acid identity to the purinoreceptor P2Y₁ and the platelet-activating-factor receptor and 28% identity to the BLT receptor¹⁶. This receptor, designated HG55, encodes a protein of 337 amino acids with a calculated molecular mass of 38,549 (Fig. 1). A patent filing (EP0874047A2), published on 28 October 1998 (ref. 17), described a complementary DNA clone (HMTMF81) that was identical to HG55. The patent claim stated that HMTMF81 could be activated by LTD₄ and LTC₄ but no data were provided. Here we describe the characterization of the human orphan GPCR HG55 as the CysLT₁ receptor.

The CysLT₁ receptor is known to signal through elevation of intracellular calcium¹⁸. We used two different *Xenopus laevis* functional systems to characterize HG55 (CysLT₁)-receptor activation (Fig. 2). In oocytes injected with HG55 (CysLT₁) cRNA, LTD₄ evoked a robust calcium-dependent chloride conductance and the response was attenuated upon repeated administration of agonist (Fig. 2a). Oocytes injected with saline (or with other orphan GPCR

MetAspGluThrGlyAsnLeuThrValSerSerAlaThrCysHisAspThrIleAspAspPheArgAsnGlnValTyrSerThrLeuTyrSerMetIleS	34
ATGGATGAACAGGAAATCTGACAGTATCTTCTGCCACATGCCATGACACTATTGATGACTTCCGCAATCAAGTGTATTCACCTTGTACTCTATGATCT	100
II	
erValValGlyPhePheGlyAsnGlyPheValLeuTyrValIleLeuLysThrTyrHisLysLysSerAlaPheGlnValTyrMetIleAsnLeuAlaVa	67
CTGTGTAGGCTTCTTTGGCAATGGCTTGTGCTCTATGCTCTATAAAACCTATCACAAGAAGTCAGCCTTCCAAGTATACATGATTAATTTAGCAGT	200
IAlaAspLeuLeuCysValCysThrLeuProLeuArgValValTyrTyrValHisLysGlyIleTrpLeuPheGlyAspPheLeuCysArgLeuSerThr	100
AGCAGATCTACTTTGTGTGTGACACTGCTCTCGTGTGCTTATTATGTTTCAAAAGGCATTTGGCTCTTTGTGACTTCTTGTGCCGCTCAGCACC	300
III	
TyrAlaLeuTyrValAsnLeuTyrCysSerIlePhePheMetThrAlaMetSerPhePheArgCysIleAlaIleValPheProValGlnAsnIleAsnL	134
TATGCTTTGTATGTCAACCTCTATTGTAGCATCTCTTTATGACAGCATGAGCTTTTCCGGTGCATTGCAATTGTTTTCCAGTCCAGAACATTAAT	400
IV	
euValThrGlnLysLysAlaArgPheValCysValGlyIleTrpIlePheValIleLeuThrSerSerProPheLeuMetAlaLysProGlnLysAspGl	167
TGGTTACACAGAAAAAGCAGGTTTGTGTGTAGGTATTGGATTTTGTGATTTTGACAGTCTCCATTCTAATGGCCAAACCAAAAAGATGA	500
V	
uLysAsnAsnThrLysCysPheGluProProGlnAspAsnGlnThrLysAsnHisValLeuValLeuHisTyrValSerLeuPheValGlyPheIleIle	200
GAAAAATAATACCAAGTGCTTTGAGCCCCACAAGACAATCAAACTAAAAATCATGTTTGGTCTTGCAATTATGTGTCATTGTTGTGGCTTTATCATC	600
VI	
ProPheValIleIleIleValCysTyrThrMetIleIleLeuThrLeuLeuLysLysSerMetLysLysAsnLeuSerSerHisLysLysAlaIleGlyM	234
CCTTTTGTATTATAATTGTCTGTACCAATGATCATTTTGACCTTACTAAAAAATCAATGAAAAAATCTGTCAAGTCAATAAAAGGCTATAGGAA	700
VII	
stIleMetValValThrAlaAlaPheLeuValSerPheMetProThrHisIleGlnArgThrIleHisLeuHisPheLeuHisAsnGluThrLysProCy	267
TGATCATGGTGTGACCGCTGCCTTTTAGTCAGTTTCATGCCATATCATATTCAAGTACCATTCACCTTCATTTTACACAATGAACATAAACCTG	800
VIII	
sAspSerValLeuArgMetGlnLysSerValValIleThrLeuSerLeuAlaAlaSerAsnCysCysPheAspProLeuLeuTyrPhePheSerGlyGly	300
TGATTTCTGTCTTAGAATGCAAGTCCGTGGTCATAACCTGTCTCTGGCTGCATCAATTGTTGCTTTGACCTCTCCTATATTTCTTTCTGGGGT	900
IX	
AsnPheArgLysArgLeuSerThrPheArgLysHisSerLeuSerSerValThrTyrValProArgLysLysAlaSerLeuProGluLysGlyGluGluI	334
AACCTTAGGAAAGGCTGTCTACATTTAGAAAGCATCTTTGTCCAGCGTGACTTATGTACCCAGAAAGAGGCTCTTTGCCAGAAAAGGAGAAGAA	1000
leCysLysValEnd	337
TATGTAAAGTATAG	1014

Figure 1 Nucleotide and deduced amino-acid sequences of the HG55 (CysLT₁) receptor. The putative transmembrane segments I–VII (based on the hydrophobicity profile) are underlined, asterisks (*) denote potential N-glycosylation

sites, crosses (†) denote potential protein kinase C phosphorylation sites, and hash symbols (#) denote potential protein kinase A phosphorylation sites. GenBank accession number, AF119711.

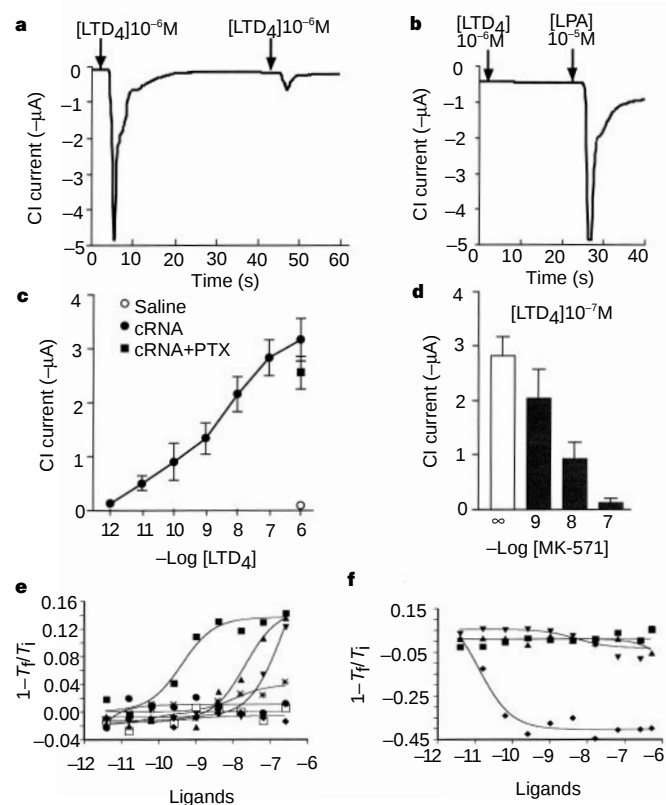


Figure 2 Functional activation of HG55 (CysLT₁) receptor in *X. laevis* cells. **a-d**, Oocyte calcium-dependent chloride conductance and **e, f**, melanophore pigment aggregation. Responses of oocytes injected with HG55 (CysLT₁) cRNA (**a, c**) or saline (**b**) and challenged with LTD₄ or LPA. **c**, Responses elicited after challenge with LTD₄ of saline-injected oocytes (open circles; $n = 16$ oocytes), HG55 (CysLT₁)-cRNA-injected oocytes (filled circles; $n = 10-16$ oocytes per point $\pm$ s.e.), and HG55 (CysLT₁)-cRNA-injected oocytes pretreated with pertussis toxin (filled squares; 100 ng/ml⁻¹, 24 h, $n = 10$ oocytes). **d**, Responses from HG55 (CysLT₁)-cRNA-injected oocytes ($n = 10-16$ oocytes) to LTD₄ pretreated with MK-571. **e**, Pigment dispersion responses in melanophores transfected with HG55 (CysLT₁) in response to LTD₄ (filled squares; EC₅₀ = 0.4 nM), LTC₄ (filled upright triangles; EC₅₀ = 21 nM), LTE₄ (filled inverted triangles; EC₅₀ = 212 nM), LTB₄ (filled diamonds), 5(S)-HETE (filled circles), LTD₄ plus 500 nM MK-571 (open squares) and control pcDNA3.1-transfected melanophores plus LTD₄ (asterisks). **f**, Pigment aggregation in melanophores transfected with HG55 (CysLT₁) and challenged with LTD₄ (filled squares), transfected with pcDNA3.1 and challenged with LTB₄ (filled upright triangles), or transfected with BLT and challenged with LTD₄ (filled inverted triangles) and LTB₄ (filled diamonds; EC₅₀ = 0.01 nM).

cRNAs; data not shown) did not respond to LTD₄ but were able to respond to lysophosphatidic acid (LPA), a ligand for an endogenous oocyte GPCR (Fig. 2b). Dose-dependent LTD₄ stimulation of oocyte calcium-activated chloride conductance was observed with a half-maximal effective concentration (EC₅₀) of 3.0 nM $\pm$ 0.2 nM (mean $\pm$ standard deviation) and a threshold response of about 10⁻¹¹ M (Fig. 2c). Pretreatment with pertussis toxin (100 ng ml⁻¹) resulted in a $\sim$ 10% lower maximal response to LTD₄, indicating limited signalling of HG55 (CysLT₁) through G_{iα/o}-linked pathways in the oocyte (Fig. 2c). Pretreatment of HG55 (CysLT₁) oocytes with MK-571, a selective CysLT₁-receptor antagonist¹⁹, blocked subsequent responses to 10⁻⁷ M LTD₄ in a dose-dependent manner (Fig. 2d). *X. laevis* melanophores respond to agonist-activated endogenous GPCRs by intracellular dispersion (as a result of intracellular calcium or cyclic AMP elevation) or aggregation (owing to cAMP lowering) of their pigment melanosomes²⁰. This system can be appropriated by human receptors that are transiently

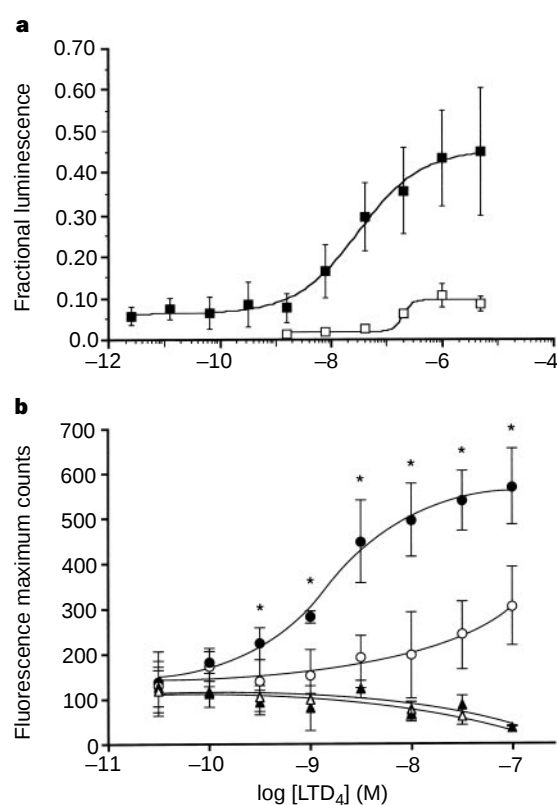


Figure 3 Functional expression of HG55 (CysLT₁) receptor in Cos-7 cells. **a**, Aequorin assay²¹. Cos-7 cells were transiently transfected with the aequorin expression vector, AEQ-pCDM and HG55 (CysLT₁)-pCEP4 or control (pCEP4) plasmids directly in 96-well plates, charged with the aequorin cofactor, coelenterazine cp, and challenged with LTD₄. Fractional luminescence responses of the HG55 (CysLT₁)-receptor (filled squares; $n = 3$) and control (open squares; $n = 2$) transfected cells to LTD₄ are plotted as a function of the ligand concentration. Data points are the means of duplicates for each of the separate experiments $\pm$ s.e. **b**, FLIPR assay. Cos-7 cells transiently transfected with HG55 (CysLT₁)-receptor cDNA or vector alone (control) were challenged with LTD₄, and their response was analysed in FLIPR (see Methods). Responses of cells transfected with the HG55 (CysLT₁) receptor to LTD₄ (by molar concentration) in the presence (open circles) or absence (filled circles) of 100 nM MK-571; responses of control cells to LTD₄ (by molar concentration) in the presence (open triangles) or absence (filled triangles) of 100 nM MK-571. Response of vehicle-treated cells was subtracted from response of agonist-treated cells. Data points are the means of duplicates for each of three separate experiments $\pm$ s.e. The asterisks indicate a probability of $P \leq 0.05$ that the response of cells transfected with the HG55 (CysLT₁) receptor and the control cell response are the same.

transfected into these cells²¹. LTD₄, LTC₄ and (very weakly) LTE₄, but not LTB₄ or 5(S)-HETE (hydroxyeicosatetraenoic acid) caused dose-dependent activation of pigment dispersion in melanophores transfected with the cDNA from the HG55 (CysLT₁) receptor. (Fig. 2e). LTD₄ was the most potent agonist in this system with an EC₅₀ value of 0.4 nM. Pretreatment of melanophores expressing HG55 (CysLT₁) receptor with MK-571 (0.5 μM) blocked LTD₄ activation (Fig. 2e). Control melanophores showed weak dispersion responses to >10 nM LTD₄ (Fig. 2e). To observe G_{iα} coupling, melanophores were preset to a pigment-dispersed state²¹. In this assay melanophores transfected with HG55 (CysLT₁)-receptor cDNA showed no aggregation response to LTD₄ (Fig. 2f). In contrast, melanophores transfected with the cDNA for the human BLT receptor, used as a positive control of a G_{iα}-coupled receptor, showed dose-dependent aggregation to LTB₄ with an EC₅₀ of 0.01 nM and no response to LTD₄ (Fig. 2f).

The ability of LTD₄ to activate the HG55 (CysLT₁) receptor that

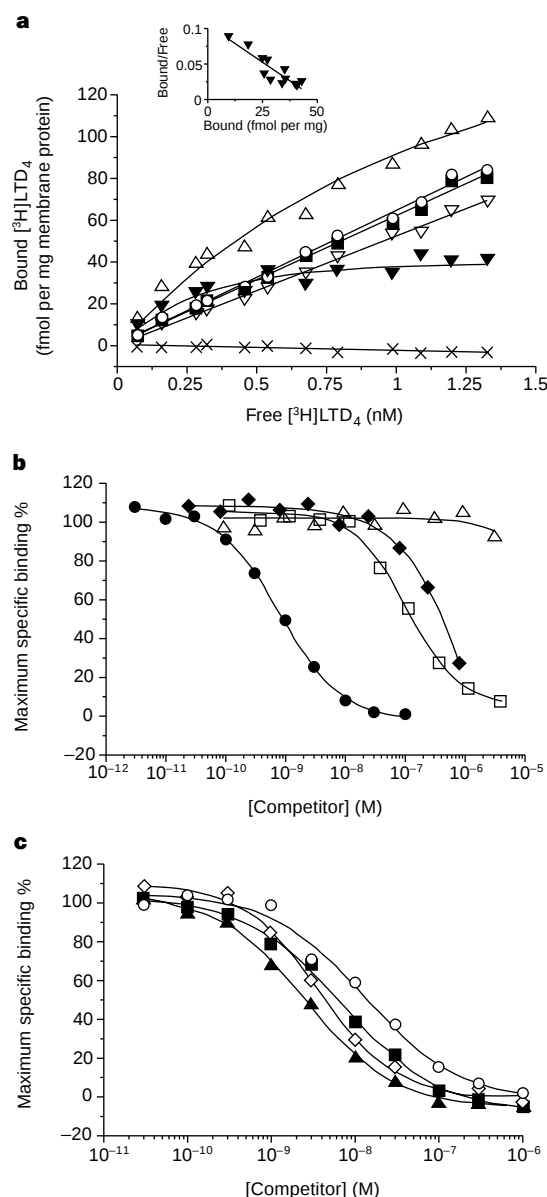


Figure 4 [^3H]LTD $_4$ -specific binding to Cos-7 cell membranes expressing the HG55 (CysLT $_1$) receptor. **a**, Saturation analysis of [^3H]LTD $_4$ -specific binding to membranes from Cos-7 cells expressing HG55 (CysLT $_1$) receptor (open upright triangles, total; open inverted triangles, non-specific; filled inverted triangles, specific) or vector alone (control) (filled squares, total; open circles, non-specific; crosses, specific). Inset, Scatchard representation of the specific binding isotherm from Cos-7 cells membranes expressing HG55 (CysLT $_1$). **b**, Competition for [^3H]LTD $_4$ -specific binding to Cos-7 cell membranes expressing HG55 (CysLT $_1$) with leukotrienes and **c**, CysLT $_1$ -selective antagonists. Representative titration curves from the same experiment are shown where at least two independent titration experiments, done in duplicate, were performed. LTD $_4$ (filled circles), LTE $_4$ (open squares), LTC $_4$ (filled diamonds), LTB $_4$ (open triangles), MK-571 (open circles), pranlukast (filled squares), montelukast (open diamonds) and zafirlukast (filled triangles).

was expressed transiently in Cos-7 monkey kidney cells was determined using both aequorin luminescence and Calcium Green-1 fluorescence as a measure of intracellular calcium mobilization. Calcium flux in Cos-7 cells transfected with HG55 (CysLT $_1$)-receptor cDNA was activated in a dose-dependent manner by LTD $_4$ with EC $_{50}$ values of ~ 2 – 35 nM (Fig. 3). LTC $_4$ was a full agonist but was around ten times less potent than LTD $_4$, whereas LTE $_4$ behaved as a weak agonist and there was no activation by LTB $_4$

Table 1 Affinities of leukotrienes and CysLT $_1$ antagonists for the HG55 (CysLT $_1$) receptor

	IC $_{50}$ (nM)	
	(–) HSA	(+) HSA
LTB $_4$	>3000, >3000	ND
LTC $_4$	360, 346	ND
LTD $_4$	0.9 $\pm$ 0.1	ND
LTE $_4$	274, 107	ND
MK-571	10.4 $\pm$ 1.6	20, 24
Montelukast	2.3, 4.5	1.2, 2.7
Pranlukast	4.3, 7.2	54, 91
Zafirlukast	1.9, 2.6	3.9, 6.7

Radioligand binding assays were performed in the absence and presence of 0.05% (w/v) HSA. Results are from two independent experiments with the individual values shown, except for LTD $_4$ and MK-571 (no HSA) where values are mean $\pm$ s.e.; $n = 3$. ND, not determined.

(data not shown). Conversely, control cells showed either a weak or no endogenous response to LTD $_4$ and LTC $_4$ (Fig. 3, and data not shown). Preincubation of these cells with 100 nM MK-571 significantly blocked the LTD $_4$ activation in the fluorescence assay (Fig. 3b) and in the aequorin assay (data not shown).

We next performed saturation analysis of [^3H]LTD $_4$ binding to membranes prepared from Cos-7 cells, transfected with either HG55 (CysLT $_1$)-receptor cDNA or vector alone (control) (Fig. 4a). A single population of high-affinity [^3H]LTD $_4$ -specific binding sites was detected in membranes from cells expressing the HG55 (CysLT $_1$) receptor, with an equilibrium dissociation constant (K_d) of 0.3 nM and maximum number of specific binding sites (B_{max}) of ~ 50 fmol per mg membrane protein (Fig. 4a), which is similar to that reported for human lung membranes 22,23 . There was no detectable [^3H]LTD $_4$ -specific binding in membranes prepared from control cells over the range of radioligand concentrations used for saturation analysis (Fig. 4a). Cysteinyl leukotrienes and CysLT $_1$ antagonists competed for [^3H]LTD $_4$ -specific binding to membranes prepared from Cos-7 cells transfected with HG55 (CysLT $_1$)-receptor cDNA (Fig. 4b, c, Table 1). The rank order of affinities for the leukotrienes was as expected, with LTD $_4$ $\gg$ LTC $_4$ $\gg$ LTB $_4$. (Fig. 4b, Table 1). The affinity of LTE $_4$ for the HG55 (CysLT $_1$) receptor was approximately 200-fold lower than that of LTD $_4$. This is comparable to the ratios for LTD $_4$ to LTC $_4$ observed in competition-binding studies using membranes prepared from human U937 cells (136-fold) and human lung (66-fold) 22 . In contrast, in membrane preparations from guinea pig lung, LTE $_4$ is only tenfold less active than LTD $_4$ (ref. 22). The affinity of LTC $_4$ for the HG55 (CysLT $_1$) receptor was ~ 350 -fold lower than that of LTD $_4$. This is comparable to the ratios for LTD $_4$ to LTC $_4$ in competition-binding studies using membranes from U937 cells (345-fold) and human lung (120-fold) 22 . The CysLT $_1$ antagonists MK-571, montelukast (Singulair), zafirlukast (Accolate) and pranlukast (Onon), all displayed high affinity for the HG55 (CysLT $_1$) receptor (Fig. 4c, Table 1). The latter three compounds are examples of structurally distinct, selective CysLT $_1$ -receptor antagonists that are marketed for the treatment of asthma. The CysLT $_1$ -receptor antagonists were also evaluated in the presence of 0.05% (w/v) human serum albumin (HSA; Table 1). HSA is routinely added to receptor assays in drug discovery programmes to assess the potential for serum protein binding. A lack of protein binding may be critical for the clinical efficacy of a given compound. Under these conditions the affinities of montelukast and zafirlukast for HG55 (CysLT $_1$) did not change significantly, but the half-maximal inhibitory concentration (IC $_{50}$) values for MK-571 and, in particular, pranlukast increased.

Northern blot analysis of various human tissues for HG55 (CysLT $_1$)-receptor mRNA showed a weakly expressed transcript of ~ 3 kilobases (kb). The highest mRNA expression was observed in spleen and peripheral blood leukocytes, with lower expression in several other tissues including lung (Fig. 5a). Given the crucial importance of cysteinyl leukotrienes in human asthma we investi-

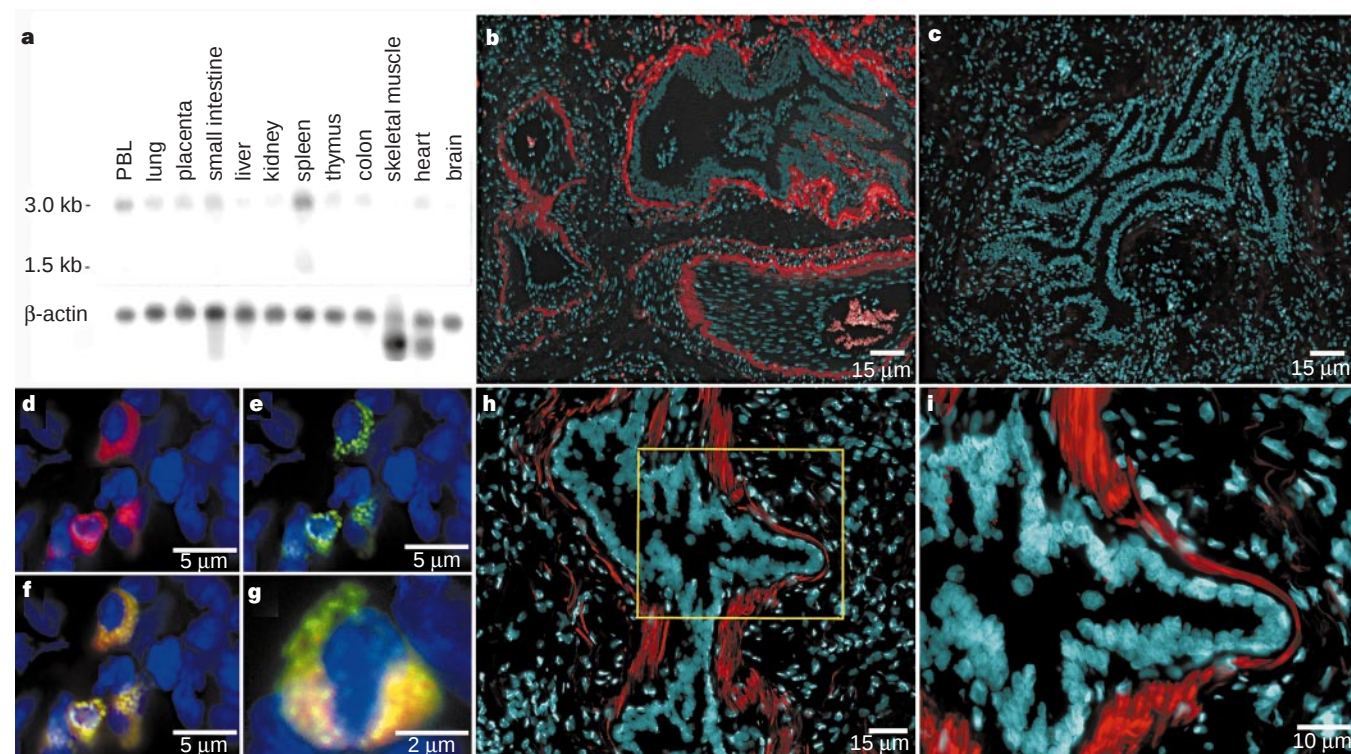


Figure 5 Expression studies of the CysLT₁ receptor. **a** Human tissue northern blot; **b–i**, *in situ* hybridization in human lung. **b**, **h–i**, Antisense probe, showing expression in bronchial smooth muscle; box in **h** is magnified in **i**; **c**, sense control probe. **d–g**, Expression of CysLT₁ receptor in human lung macrophages: **d**, CysLT₁ receptor antisense *in situ*; **e**, Mac1 immunohistochemistry; **f–g**, combined

CysLT₁ receptor *in situ* and Mac1 immunohistochemistry, demonstrating co-localization. Similar results were seen using the D11 antibody (data not shown). Cellular nuclei are seen in teal pseudocolour (**b–c**, **h–i**), or conventional DAPI blue (**d–g**); CysLT₁ receptor *in situ* signal as Texas Red fluorescence; Mac1 immunoreactivity as FITC green fluorescence; and CysLT₁/Mac1 co-localization as yellow.

gated the distribution of the HG55 (CysLT₁) receptor in human lung by *in situ* hybridization. HG55 (CysLT₁)-receptor mRNA expression was seen in the peribronchial and peribronchiolar smooth muscle cells at all levels of the respiratory tree (data not shown), including the segmental bronchi and bronchioles (Fig. 5b, h, i). No signal was detected using the sense control (Fig. 5c). At higher magnifications of a human bronchiole (Fig. 5h, i), restricted distribution of HG55 (CysLT₁)-receptor mRNA was seen in smooth muscle bundles, with minimal receptor mRNA detected in epithelial cells lining the airway lumen. HG55 (CysLT₁)-receptor mRNA was also demonstrated in some lung macrophages as identified by morphology, location and antibody co-localization with D11 and Mac1 antisera (Fig. 5d–g, and data not shown). The HG55 (CysLT₁) gene was mapped by monochromosomal somatic-cell-hybrid and radiation-hybrid analyses to the long arm of the human X chromosome at bands 13–21 (Xq13–Xq21). Until now, no asthmatic disease markers have been shown to map to this region.

In conclusion, we have cloned and characterized the human CysLT₁ receptor for cysteinyl leukotrienes. All three marketed cysteinyl leukotriene antagonists, namely montelukast (Singulair), zafirlukast (Accolate) and pranlukast (Onon) are potent competitors for the binding of radiolabelled LTD₄ to the CysLT₁ receptor. The localization of the CysLT₁ receptor in smooth muscle and alveolar macrophages correlates well with the antibronchoconstrictive and anti-inflammatory nature of these new therapeutic entities. Expression studies of this receptor in asthmatic and allergic tissues will further our understanding of the pathobiology of the cysteinyl leukotrienes. □

Methods

Cloning of HG55 cDNA. HG55 cDNA was cloned using the polymerase chain reaction (PCR) to amplify the translational open reading frame (ORF) from the EST cDNA (id_1357580) using the oligonucleotides 5'-ACGGTACCATGGAT-GAAACAGG-3' (sense) and 5'-CCTCTAGAAATGGGTTTAACTATAC-3'

(antisense). This amplified product was subcloned into the Acc65I and XbaI sites of modified pcDNA3.1+ (Invitrogen) and the sequence of the inserted DNA was determined.

***X. laevis* oocyte expression.** HG55 (CysLT₁) cRNA was prepared *in vitro* (T7 mMessage mMachine, Ambion) using T7 RNA polymerase in the presence of a capping analogue. Between 20 and 40 ng of the capped cRNA was injected into stage V–VI *X. laevis* oocytes as described previously^{24,25}. Compounds were delivered as a 30-μl aliquot over a period of 1–2 s and the recording chamber was washed with 30% methanol between oocyte assays to remove residual ligand. *X. laevis* were purchased from Xenopus; LTD₄, LTC₄ from Cayman; LTB₄, LTE₄, MK-571 from BIOMOL; and 1-oleoyl LPA from Avanti Polar Lipids.

Melanophore functional assay. Growth of *X. laevis* melanophores and fibroblasts, and DNA transfections by electroporation, were done as described^{20,26}. Absorbance readings at 600 nm were measured using a Bio-Tek Elx800 Microplate reader (ESBE Scientific) before (*T*_i) and after (*T*_f) incubation for 1.5 h with the ligands. The BLT-receptor cDNA was a gift from T. Shimizu, University of Tokyo¹⁴.

Aequorin luminescence functional assay. Cos-7 cells were transiently transfected using Lipofectamine (GIBCO BRL), according to the manufacturer's instructions, with HG55 (CysLT₁)-pcDNA3.1 or pcDNA3.1 and AEQ-pCDM plasmids, 30 ng each per well of a 96-well culture plate and used in the aequorin luminescence assay as described²⁷.

Fluorescence calcium-imaging functional assay. Cos-7 cells were transiently transfected with HG55 (CysLT₁)-receptor DNA for 48 h, detached and replated into 96-well plates at a density of 2×10^4 cells per well. After 24 h cells were loaded with Calcium Green-1 in the presence of 2.5 mM probenecid. After washing, the cells were pretreated for 5 min with 100 nM MK-571 or 0.1% dimethylsulphoxide (DMSO) vehicle, and then with varying concentrations of LTD₄. Fluorescence output was measured in a Molecular Devices Fluorometric Imaging Plate reader (FLIPR).

Radioligand binding assays. [³H]LTD₄ (146 Ci mmol⁻¹) was from NEN Life Science Products; LTB₄, LTD₄, LTE₄, MK-571, montelukast, pranlukast and zafirlukast were synthesized by the Department of Medicinal Chemistry at

Merck Frosst; and LTC₄ was from Cayman. Cos-7 cell transfection, harvesting and membrane preparation were done as reported²⁸.

Tritiated LTD₄-binding assays were done as described²² with the exception that the reaction was initiated by addition of 250 µg of membrane protein. LTD₄-specific binding was calculated by subtracting non-specific binding, determined in the presence of 10 µM MK-571, from total binding (that was less than 10% of the radioligand concentration in the incubation). Specific binding accounted for 55–65% of the total binding and was linear with respect to the concentrations of radioligand and protein present in the incubation. (Average total and non-specific binding of [³H]LTD₄ for 250 µg of HG55 (CysLT₁)-receptor membrane protein was 3,000 d.p.m. and 1,100 d.p.m., respectively). Saturation analysis data were analysed by nonlinear transformation of the calculated [³H]LTD₄-specific binding saturation isotherm using GraphPad Prism software to generate K_d and B_{max} values.

Northern blot analysis. Human northern blots (Clontech) were hybridized with a HG55 (CysLT₁) coding-region cDNA and a β-actin control cDNA labelled using random priming with [³²P]-dCTP and exposed for either 3 h (β-actin) or 18 h (CysLT₁).

In situ hybridization analysis and immunohistochemistry. The oligonucleotide probes used for the *in situ* studies were 5'-GAACATAATAGACCACACGAGAGGCAGTG-3' (antisense) and 5'-CACTGCCTCTCCGTGTGGTCTATTATGTTTC-3' (sense). Tailing of oligonucleotides with biotin-16-dUTP (Boehringer Mannheim) was done as described by the manufacturer except for replacement of digoxigenin-dUTP with biotin-16-dUTP in the tailing reaction. *In situ* hybridization was done on 6-µm cryostat sections of surgical human lung specimens (National Disease Research Interchange) using 2 pmol ml⁻¹ of labelled oligonucleotides for 18 h at 37 °C. Bound probe was visualized using Tyramide Signal Amplification (NEN Life Sciences) and Texas Red fluorescence detection according to the manufacturer's instructions. Sections were counterstained with 4,6-diamidino-2-phenylindole (DAPI; Molecular Probes) to visualize cell nuclei, and images were digitally acquired and reassembled using a MicroMax CCD camera (Princeton Instruments) and Metamorph imaging program (Universal Imaging). In Fig. 5b, c, h, i cell nuclei are stained with teal blue pseudocolour, whereas in Fig. 5d–g cell nuclei are stained with conventional DAPI blue. Immunohistochemistry was done after *in situ* hybridization with antibodies D11 (BioGenex) or Mac1 (Boehringer M1/70H1) at a dilution of 5 µg ml⁻¹ for 2 h at room temperature, and products were detected with 1:300 fluorescein isothiocyanate (FITC) donkey anti-rat (immunoglobulin-γ (IgG) (Jackson ImmunoResearch). Co-localization of red (*in situ*) and green (antibody) staining is seen as yellow fluorescence.

Chromosomal localization. Two pairs of primers, a forward primer 5'-ATGGATGAAACAGGAATCTG-3' and two reverse primers 5'-TAATGCAAGACCAAAACATGA-3' and 5'-TTTATGAGGACATAGAGCACA-3', specific for HG55 5' coding sequences, were used against a monochromosomal somatic-cell-hybrid panel (Quantum) to locate the gene to chromosome X. The radiation-hybrid GeneBridge4 (Research Genetics) panel was also used and data were submitted to the Sanger Center for linkage analysis at <http://www.sanger.ac.uk/RHserver/RHserver.html>. HG55 was found to link to marker AFM308xb9 (with a likelihood score of 7.048). The Stanford G3 RH panel was used to confirm this linkage.

Received 19 March; accepted 20 April 1999.

- Dahlen, S. E., Hedqvist, P., Hammarstrom, S. & Samuelsson, B. Leukotrienes are potent constrictors of human bronchi. *Nature* **288**, 484–486 (1980).
- Piper, P. J. Formation and actions of leukotrienes. *Physiol. Rev.* **64**, 744–761 (1984).
- Lewis, R. A., Austen, K. F. & Soberman, R. J. Leukotrienes and other products of the 5-lipoxygenase pathway. Biochemistry and relation to pathology in human diseases. *N. Engl. J. Med.* **323**, 645–655 (1990).
- Coleman, R. A. et al. Prostanoid and leukotriene receptors. A progress report from the IUPHAR working parties on classification and nomenclature. *Adv. Prostaglandins Thromboxane Leukotriene Res.* **23**, 283–285 (1995).
- Charette, L. & Jones, T. J. Effects of L-serine borate on antagonism of leukotriene C₄-induced contractions of guinea pig trachea. *Br. J. Pharmacol.* **91**, 179–188 (1987).
- Labat, C. et al. A second cysteinyl leukotriene receptor in human lung. *J. Pharmacol. Exp. Ther.* **263**, 800–805 (1992).
- Jones, T. R. et al. Pharmacology of montelukast sodium (SINGULAIR™), a potent and selective leukotriene D₄ receptor antagonist. *Can. J. Physiol. Pharmacol.* **73**, 191–201 (1995).
- Reiss, T. F. et al. Montelukast, a once-daily leukotriene receptor antagonist, in the treatment of chronic asthma: a multicenter, randomized, double-blind trial. Montelukast Clinical Research Study Group. *Arch. Intern. Med.* **158**, 1213–1220 (1998).
- Knorr, B. et al. Montelukast for chronic asthma in 6- to 14-year-old children: a randomized, double-blind trial. Pediatric Montelukast Study Group. *J. Am. Med. Assoc.* **279**, 1181–1186 (1998).
- Leff, J. A. et al. Montelukast, a leukotriene-receptor antagonist, for the treatment of mild asthma and exercise-induced bronchoconstriction. *N. Engl. J. Med.* **339**, 147–152 (1998).

- Suissa, S. et al. Effectiveness of the leukotriene receptor antagonist zafirlukast for mild-moderate asthma. *Ann. Int. Med.* **126**, 177–183 (1997).
- Grossman, J. et al. Results of the first U.S. double-blind, placebo-controlled multicenter clinical study in asthma with pranlukast, a novel leukotriene receptor antagonist. *J. Asthma* **34**, 321–328 (1997).
- Metters, K. M. & Zamboni, R. J. Photoaffinity labeling of the leukotriene D₄ receptor in guinea pig lung. *J. Biol. Chem.* **268**, 6487–6495 (1993).
- Yokomizo, T., Izumi, T., Change, K., Takuwa, Y. & Shimizu, T. A G-protein-coupled receptor for leukotriene B₄ that mediates chemotaxis. *Nature* **387**, 620–624 (1997).
- Martin, V. et al. Leukotriene binding, signaling and analysis of HIV coreceptor function in mouse and human leukotriene B₄ receptor transfected cells. *J. Biol. Chem.* **274**, 8597–8603 (1999).
- Pearson, W. R. & Lipman, D. J. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448 (1988).
- Ellis, C. et al. EP874047A2: cDNA clone HMTMF81 that encodes a novel human 7-transmembrane receptor. Eur. Patent Applic. EP0874047A2 (1998).
- Chan, C. et al. Leukotriene D₄-induced increases in cytosolic calcium in THP-1 cells: dependence on extracellular calcium and inhibition with selective leukotriene D₄ receptor antagonists. *J. Pharmacol. Exp. Ther.* **269**, 891–896 (1994).
- Jones, T. R. et al. Pharmacology of L-660,711 (MK-571): a novel, potent and selective leukotriene D₄ receptor antagonist. *Can. J. Physiol. Pharmacol.* **67**, 17–28 (1998).
- McClintock, T. S., Rising, J. P. & Lerner, M. R. Melanophore pigment dispersion responses to agonists show two patterns of sensitivity to inhibitors of cAMP-dependent protein kinase and protein kinase C. *J. Cell. Physiol.* **167**, 1–7 (1996).
- Potenza, M. N., Graminski, G. F., Schmaus, C. & Lerner, M. R. Functional expression and characterization of human D2 and dD3 dopamine receptors. *J. Neurosci.* **14**, 1463–1476 (1994).
- Frey, E. A. et al. Characterization of the leukotriene D₄ receptor in dimethylsulphoxide-differentiated U937 cells: comparison with the leukotriene D₄ receptor in human lung and guinea-pig lung. *Eur. J. Pharmacol.* **244**, 239–250 (1993).
- Metters, K. M. Leukotriene receptors. *J. Lipid Mediat.* **12**, 413–427 (1995).
- Durieux, M. E., Carlisle, S. J., Salafra, M. N. & Lynch, K. R. Endogenous responses to sphingosine-1-phosphate in *X. laevis* oocytes: similarities with lysophosphatidic acid signaling. *Am. J. Physiol.* **264**, C1360–C1364 (1993).
- Hooks, S. B. et al. Characterization of a receptor subtype-selective lysophosphatidic acid mimetic. *Mol. Pharmacol.* **53**, 188–194 (1998).
- Kolakowski, L. F. Jr et al. Molecular characterization and expression of cloned human galanin receptors GALR2 and GALR3. *J. Neurosci.* **71**, 2239–2251 (1998).
- Ungrin, M. D. et al. An automated aequorin luminescence-based functional calcium assay for G-protein coupled receptors. *Anal. Biochem.* (in the press).
- Wright, D. H., Metters, K. M., Abramovitz, M. & Ford-Hutchinson, A. W. Characterization of the recombinant prostanoid DP receptor and identification of L-644,698, a novel DP agonist. *Br. J. Pharmacol.* **123**, 1317–1324 (1998).

Acknowledgements. The authors thank M. E. Durieux (Dept. Anesthesiology, Univ. Virginia) for the use of his laboratory's oocyte recording equipment, D. Pettibone and R. Gould (Dept. Pharmacology, Merck & Co., West Point, Pennsylvania) for support to the MRL orphan GPCR project, and K. Clark for preparation of the figures. This work was supported in part by grants to K.R.L. from the National Institutes of Health. This paper is dedicated to the late P. J. Piper for her pioneering work in the pharmacological characterization of the cysteinyl leukotrienes.

Correspondence and requests for materials should be addressed to J.F.E. (e-mail: jilly_evans@merck.com). DNA and amino-acid sequences have been deposited in the GenBank/EMBL/DBJ databases under accession number AF119711.

MAP kinase and Wnt pathways converge to downregulate an HMG-domain repressor in *Caenorhabditis elegans*

Marc D. Meneghini*, Tohru Ishitani†, J. Clayton Carter*, Naoki Hisamoto†, Jun Ninomiya-Tsuji†, Christopher J. Thorpe*, Danielle R. Hamill*, Kunihiro Matsumoto† & Bruce Bowerman*

* Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403, USA

† Department of Molecular Biology, Graduate School of Science, Nagoya University, and CREST, Japan Science and Technology Corporation, Chikusa-ku, Nagoya 464-8602, Japan

The signalling protein Wnt regulates transcription factors containing high-mobility-group (HMG) domains to direct decisions on cell fate during animal development¹. In *Caenorhabditis elegans*, the HMG-domain-containing repressor POP-1 distinguishes the fates of anterior daughter cells from their posterior sisters throughout development^{2,3}, and Wnt signalling downregulates POP-1 activity in one posterior daughter cell called E (refs 2, 4, 5). Here we show that the genes *mom-4* and *lit-1* are also required to downregulate POP-1, not only in E but also in other

* Present address: Department of Pharmacology, University of Washington School of Medicine, Pharmacology Box 357370, Seattle, Washington 98195-7370, USA.

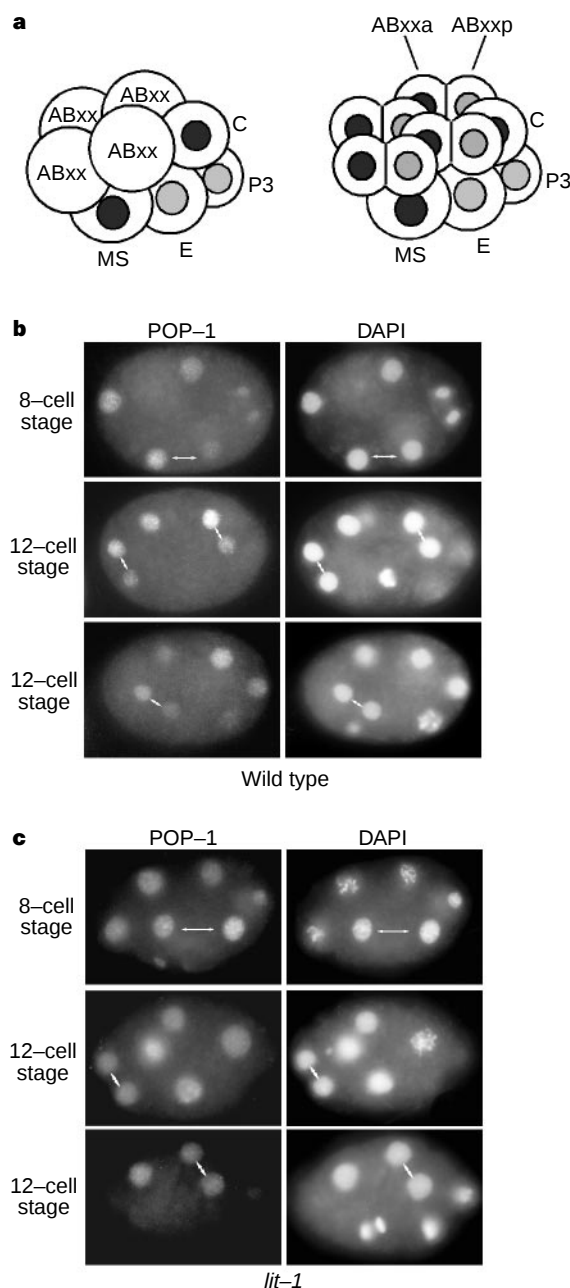


Figure 1 POP-1 asymmetries in wild-type and *lit-1* mutant embryos. **a**, Schematic representation of 8-cell stage (left) and 12-cell stage (right) *C. elegans* embryos. Shading inside each cell indicates relative levels of detectable POP-1 protein. The four cells labelled ABxx at the 8-cell stage are the parents of the four ABxxa/p sister pairs shown in the 12-cell stage embryo. **b**, **c**, Immunofluorescence micrographs showing POP-1 (left panels) and DAPI (right panels) staining to visualize nuclei in wild-type (**b**) and *lit-1* mutant (**c**) embryos. Embryos are oriented with anterior to the left and sister-cell pairs are indicated by double-headed arrows.

Table 1 POP-1 staining results

Genotype	MS/E	ABxxa/p
Wild type	0/8	0/16
<i>mom-1(or10)</i>	11/15	7/10
<i>mom-2(or42)</i>	13/18	7/12
<i>lit-1(or131)</i>	20/20	13/13
<i>mom-4(or39)</i>	9/19	13/13

Numbers of embryos showing a loss of POP-1 asymmetry. For MS/E and ABxxa/p, the fraction indicates the number of embryos with equal levels of detectable POP-1 in both anterior and posterior daughter cells/total number analysed. Embryos not showing a disruption of POP-1 asymmetry had higher levels of POP-1 in the anterior cells MS and ABxxa compared with their posterior sisters E and ABxpx, as in the wild type.

posterior daughter cells. Consistent with action in a common pathway, *mom-4* and *lit-1* exhibit similar mutant phenotypes and encode components of the mitogen-activated protein kinase (MAPK) pathway that are homologous to vertebrate transforming-growth-factor- β -activated kinase (TAK1) and NEMO-like kinase (NLK), respectively. Furthermore, MOM-4 and TAK1 bind related proteins that promote their kinase activities. We conclude that a MAPK-related pathway cooperates with Wnt signal transduction to downregulate POP-1 activity. These functions are likely to be conserved in vertebrates, as TAK1 and NLK can downregulate HMG-domain-containing proteins related to POP-1 (ref. 6).

In a four-cell-stage *C. elegans* embryo, Wnt signalling polarizes an embryonic cell called EMS such that the HMG-domain-containing protein POP-1 is detectable at high levels in its anterior daughter, MS, and at lower levels in its posterior daughter, E (refs 2, 4, 5, 7, 8; Fig. 1a). Normally, E and MS produce endoderm and mesoderm, respectively. Downregulation of POP-1 in E allows the specification of endoderm, as both E and MS produce endoderm in embryos lacking *pop-1* function². In embryos lacking the function of the Wnt pathway components MOM-1/porcupine, MOM-2/wingless or WRM-1/ β -catenin, both EMS daughters accumulate high levels of POP-1 and adopt MS-like fates^{4,5}. In other systems, Wnt signalling stabilizes cytoplasmic β -catenin which translocates to the nucleus and converts HMG-domain-containing repressors such as T-cell factor (TCF) and lymphoid enhancer factor (LEF-1) into transcriptional activators⁹⁻¹¹. In early *C. elegans* embryos, however, WRM-1/ β -catenin is required to downregulate POP-1 in E, blocking its function as a repressor rather than converting it into an activator^{4,5}. Data from *Xenopus* and *Drosophila* support both outcomes, depending perhaps on the context in which a signal is received⁹⁻¹⁵.

Because *lit-1* and *mom-4* act genetically as upstream negative regulators of *pop-1* in E (refs 5, 16), we asked whether they could be required, like Wnt signalling, to downregulate POP-1 in E (Fig. 1 and Table 1). We also examined POP-1 asymmetries in four pairs of sister cells born at the 12-cell stage which we refer to collectively as the ABxxa/p sister pairs (Fig. 1a). In wild-type embryos, both MS/E and ABxxa/p sister pairs reproducibly show POP-1 asymmetries (Fig. 1b and Table 1). In *lit-1(or131ts)* mutant embryos, we observed fully penetrant losses of MS/E and ABxxa/p POP-1 asymmetries (Fig. 1c and Table 1), consistent with their nearly fully penetrant loss of endoderm (see Methods) and with *lit-1* being required for anterior-posterior fate decisions throughout most of embryogenesis¹⁶. *mom-4* mutant embryos exhibit only partially penetrant losses of both endoderm^{4,5} and MS/E POP-1 asymmetry (Table 1). However, *mom-4* resembles *lit-1*, and the downstream Wnt component *wrm-1*/ β -catenin³, in that *mom-4* mutations cause a fully penetrant loss of ABxxa/p POP-1 asymmetries (Table 1). In contrast, mutations in the upstream Wnt pathway components *mom-1*/*porc* or *mom-2*/*wg* cause only partially penetrant losses of POP-1 asymmetry in MS/E and ABxxa/p sister-cell pairs^{4,5} (Table 1). Thus, *mom-4* and *lit-1* are required to downregulate POP-1 but are

Table 2 Tissue differentiation in wild-type, *mom-2* and *mom-4* embryos

Blastomere isolated	Number of body-wall muscle cells			
	Wild type	<i>mom-2</i>	<i>mom-4</i>	<i>lit-1(t1512)</i> [†]
Intact embryo	81 (ND)	ND	27 (10) 20-35	35
E	0 (5)	ND	6 (3) 5-6	4
MS	24 (5) 22-25	24 (6) 22-25	2 (18) 0-7	3
C	30 (5) 28-33	ND	8 (15) 0-12	9
Ca/Cp	15 (3) 15-16	16 (11) 14-20*	ND	
Blastomere isolated	Number of pharyngeal muscle cells			
	Wild type	<i>mom-2</i>	<i>mom-4</i>	<i>lit-1(t1512)</i> [†]
E	0 (3)	ND	16 (8) 13-18	15
MS	13 (6) 10-15	11 (3) 10-13	22 (8) 18-26	24

All embryonic cells except the one indicated were killed using a laser microbeam. Partial embryos were allowed to develop, then fixed and stained with tissue-specific antibodies. The number of operated embryos scored is indicated in parentheses, followed by the range. ND, not determined.

* Previously reported⁵.

[†] Ref. 16.

more similar in their requirements to WRM-1/ β -catenin than to MOM-1/porcupine or MOM-2/wingless.

To compare the *mom-4* and *lit-1* mutant phenotypes further, we examined the specification of body-wall and pharyngeal muscle made by the two embryonic cells MS and C (Table 2). Owing to serial posterior-to-anterior fate transformations in their cell lineages in *lit-1* mutants, MS and C both produce abnormally few body-wall muscle cells, and MS makes extra pharyngeal muscle cells¹⁶ (Table 2). Similarly, in *mom-4* mutants both MS and C produce less body-wall muscle than normal, and MS makes extra pharyngeal muscle (Table 2). By comparison, MS and C develop relatively normally in *mom-2/wg* mutant embryos (Table 2), further indicating that a *mom-4*–*lit-1* pathway can be distinguished from *mom-2/wg* signalling.

To determine its relationship to the Wnt pathway at a molecular level, we positionally cloned *mom-4* using germline transformation (see Methods), and we identified a unique lesion in the DNA sequences of three mutant alleles (Fig. 2a). The predicted MOM-4 protein is most similar to mammalian TAK1, a member of the MAP-kinase-kinase kinase (MAP3K) family¹⁷. In addition to their related kinase domains, MOM-4 and TAK1 share sequence similarity in a carboxy-terminal region unique to the TAK1/MOM-4 family, which

includes human, *Drosophila* and *Xenopus* homologues.

Because TAK1 kinase activity requires a binding protein called TAB1, we performed a two-hybrid screen to identify proteins that regulate MOM-4 and identified TAP-1, a TAB1-like protein (Fig. 2b). We also transfected human embryonic 293 cells with expression vectors encoding a Myc-epitope-tagged TAP-1 and a Flag-epitope-tagged MOM-4. We found that Myc–TAP-1 was present in Flag–MOM-4 immunocomplexes and that each protein appeared to become extensively phosphorylated when co-expressed (Fig. 3a), as seen with TAK1 and TAB1 (ref. 18). When over-expressed with TAB1, TAK1 can phosphorylate MAP2Ks such as MKK3, MKK4 and MKK6, and activate the JNK (c-Jun N-terminal kinase) and p38 MAPK pathways¹⁷. We therefore overexpressed Flag–MOM-4 with or without TAP-1 in 293 cells and immunoprecipitated Flag–MOM-4 from cell lysates for use in protein kinase assays with His–MKK6 as a substrate (Fig. 3b). As expected, co-expressed TAP-1 strongly enhanced the phosphorylation of MKK6 by MOM-4 *in vitro*.

We also investigated whether MOM-4 could activate JNK, by co-expressing Flag–MOM-4 with an amino-terminal haemagglutinin (HA)-epitope-tagged JNK, again with or without TAP-1. HA–JNK immunocomplexes from cell extracts were assayed using



Figure 2 *mom-4* and *tap-1* encode proteins homologous to TAK1 and TAB1, respectively. In the consensus lines, identical residues are noted by the amino-acid and similar residues by a '+' symbol. **a**, Alignment of MOM-4 and TAK1. The positions of conserved kinase subdomains I–XI are shown above the sequences. The mutation sites in *or49*, *or39* and *or11* are shown. The mutation in *or11* is an insertion of eight bases at the cDNA position corresponding to arginine 396.

Although the *mom-4(or49)* mutation causes a less penetrant endoderm defect than *or39* (ref. 5), the *or49* lesion introduces an earlier stop codon. We therefore conclude the read-through of premature stop codons is likely and do not think our *mom-4* alleles are null, perhaps explaining the leakiness of the endoderm defect in *mom-4* mutant embryos. **b**, Alignment of TAP-1 with TAB1. Overall, TAP-1 and TAB1 are 24% identical and 50% similar.

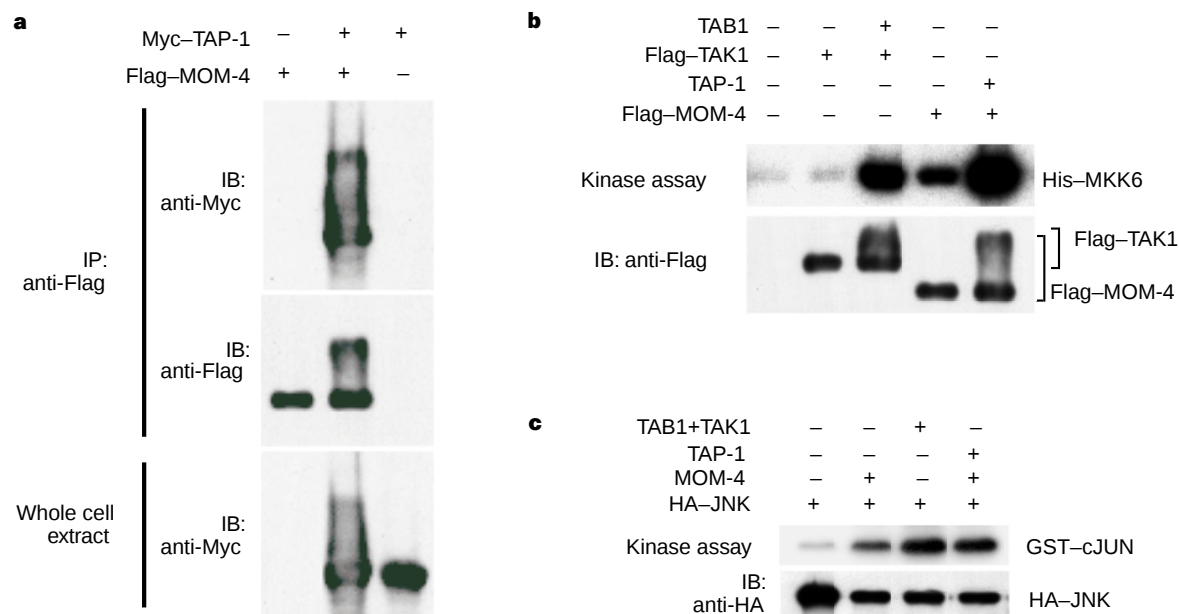


Figure 3 TAP-1 binds to MOM-4 and promotes MOM-4 MAP3K activity. **a**, Association of TAP-1 with MOM-4. Cell lysates from 293 cells transfected with expression vectors as indicated were immunoprecipitated (IP) with anti-Flag. Immunoprecipitates were immunoblotted (IB) with each antibody. Expression of Myc-TAP-1 was monitored. **b, c**, Activation of MOM-4 by TAP-1. 293 cells were

transfected with expression vectors as indicated. Flag-MOM-4 and HA-JNK were immunoprecipitated with anti-Flag (**b**) and anti-HA (**c**), respectively. The immune complexes were subjected to *in vitro* kinase assays using HIS-MKK6 (**b**) and GST-cJUN (**c**) as substrates. The immunoprecipitates were also immunoblotted with each antibody.

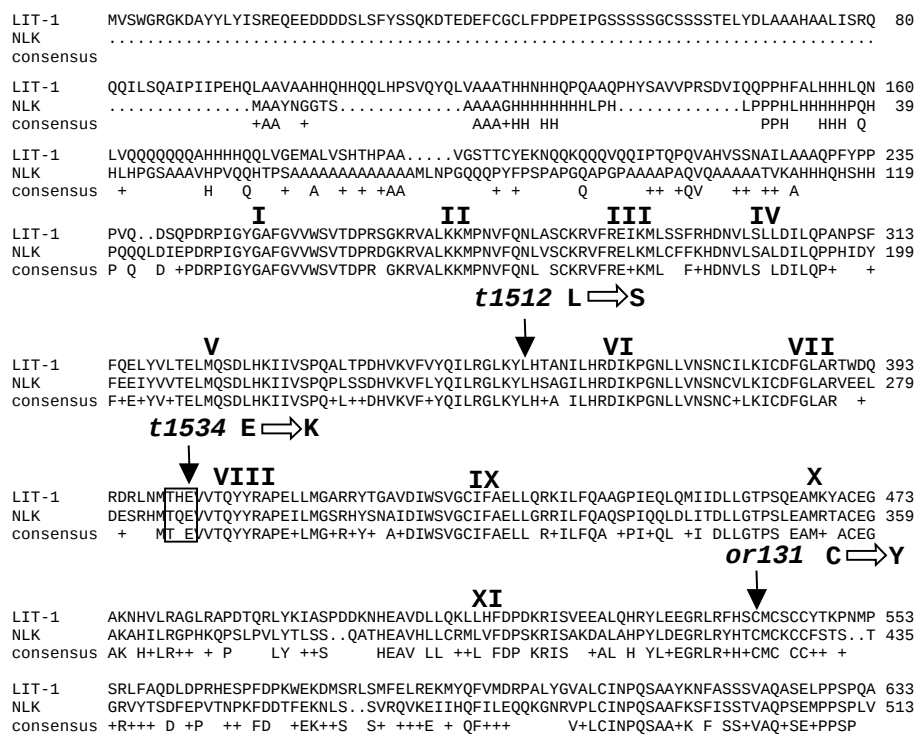


Figure 4 Alignment of LIT-1 and NLK. In the consensus line, identical residues are noted by the amino acid and similar residues by a '+' symbol. The position of

conserved kinase subdomains I-XI are shown above the sequences. The mutation sites in *t1512*, *t1534* and *or131* are shown.

glutathione-S-transferase (GST)-cJun as a test substrate (Fig. 3c). In the absence of co-transfected MOM-4, JNK exhibits basal activity. Co-expression of MOM-4 activates JNK, and TAP-1 further enhances activation. Thus, MOM-4 functions as a MAP3K, and MOM-4 and TAK1 share functionally important interactions with related binding proteins.

To assess its role in *C. elegans*, we used RNA interference (RNAi)¹⁹ to reduce *tap-1* function. RNAi of *tap-1* in a wild-type background caused no embryonic lethality. However, *tap-1* RNAi significantly

enhanced the penetrance of the endoderm defect caused by a weak *mom-4* allele, *or11*: 48% (86/176) of injected embryos lacked gut compared with 15% (28/187) of mock-injected controls. The interaction appears to be specific, as enhancement occurs only with RNA corresponding to *tap-1* or to *mom-4* itself (see Methods), further supporting the view that *tap-1* positively regulates *mom-4*.

After noting a close relative of the *Drosophila* MAP-kinase-related gene *nemo* in the sequenced *C. elegans* genome, designated W06F12.1a and located (like *lit-1*) on the far right arm of chromo-

some III, we determined that *lit-1* is W06F12.1a. Microinjecting W06F12.1a RNA causes a highly penetrant *lit-1*-like phenotype: 249 out of 252 RNAi embryos lacked endoderm. Using germline transformation, we rescued *lit-1* mutants with the cosmid W06F12 (ref. 20; and see Methods), and we also identified missense mutations in W06F12.1a in each of three temperature-sensitive *lit-1* alleles (Fig. 4).

The predicted LIT-1-kinase domain is most similar to the kinase domain of a murine relative, NLK²¹ (Fig. 4). Directly C-terminal to the kinase domain of LIT-1 is another region of sequence similarity shared only by NEMO and NLK^{21,22} and mutated in the *lit-1* allele *or131* (Fig. 4). In an accompanying report⁶, we show that introducing the same mutation into the C-terminal region of NLK eliminates its ability to bind and downregulate a vertebrate HMG-domain-containing repressor.

In kinase subdomain VIII of all known MAPKs, a conserved TXY tripartite motif is phosphorylated by upstream-activating MAP2Ks. However, LIT-1, NEMO and NLK all have TXE at this site^{21,22} (boxed in Fig. 4). Nevertheless, because LIT-1 functions in a pathway with MOM-4, a functional MAP3K, we think it is likely that LIT-1, NEMO and NLK represent a conserved subfamily of MAP kinases in which the glutamate in TXE perhaps obviates the need for tyrosine phosphorylation. We presume that MOM-4 functions upstream of LIT-1, as demonstrated for their mammalian homologues⁶. Thus, a conserved MAP-kinase-related pathway converges with Wnt signalling to downregulate an HMG-domain-containing repressor in *C. elegans*.

The vertebrate LIT-1 homologue NLK binds and phosphorylates the HMG-domain protein TCF-4, downregulating the DNA-binding activity of a TCF-4/ β -catenin complex¹⁰. Consistent with WRM-1/ β -catenin and LIT-1 both being required to downregulate POP-1, NLK may block DNA binding only when TCF-4 is in a complex with β -catenin⁶. We suspect that LIT-1 and WRM-1 similarly downregulate POP-1 DNA-binding activity. The observed POP-1 asymmetries may be a secondary effect, caused by reduced epitope accessibility, a decrease in stability or nuclear export of modified POP-1.

The apparently different outputs of Wnt signalling—downregulating HMG-domain-containing repressors versus converting them into activators—can be explained by our findings. In all cases, Wnt signalling stabilizes β -catenin. In *C. elegans* embryos, WRM-1/ β -catenin then functions with LIT-1 to downregulate POP-1. This interaction is conserved in vertebrates, although the precise developmental contexts in which it occurs are unknown, and other inputs could have similar consequences. Upstream factors that activate the MOM-4–LIT-1 pathway are unknown. Analysis of the MOM-4 binding protein TAP-1 may facilitate identification of upstream polarity cue(s) that converge with Wnt signalling to downregulate POP-1 activity. □

Methods

Strains, alleles, microscopy and laser ablations. *C. elegans* culturing and genetics were performed as described^{23,24}. N2 Bristol was the wild-type strain. Mutant strains used were: EU446 *unc-13(e1091) mom-4(or39) I/hT2* (I;III), EU230 *mom-4(or49) dpy-5(e61) I/hT2* (I;III) EU443 *unc-13(e1091) mom-4(or11) I/hT2* (I;III), EU463 *unc-13(e1091) mom-5(zu193)/mom-4(or39) lin-11(n566)*, EU308 *mom-1(or10) dpy-6 e14) X/szT1* (I;X), EU278 *mom-2(or42) V/DnT1* (IV;V), EU431 *rol-1(e91) mom-3(or78) lin-7(e1413)/mnC1* II; *him-5(e1490)* V, EU452 *unc-13(e1091) mom-5(zu193) I/hT2* (I;III), EU603 *lit-1(or131ts)* III; *him-8(e1489)*, EU551 *lit-1(or131ts)* III, GE2244 *unc-32(e189) lit-1(t1512)/qC1 dpy-19(e1259) glp-1(q339)* III; *him-3(e1147)* IV. Laser ablations, embryo fixation and antibody staining were performed as described^{3,25,26}. We used monoclonal antibodies 3NB12 and NE8 4C6.3 (refs 16, 27) and P4G4 (ref. 3) to detect pharyngeal and body-wall muscle cells, and POP-1 protein, respectively.

Cloning of *mom-4*, *lit-1* and *tap-1*. Cosmid rescue of *mom-4* and *lit-1* were achieved as described²⁰. RNAi was performed on candidate genes on these cosmids as described¹⁹. Only RNA corresponding to the gene F52F12.3

enhanced the penetrance of the endoderm defect of *mom-4(or11)*. We isolated *lit-1(or131)* in a screen for temperature-sensitive maternal-effect embryonic-lethal mutants (D.H. and B.B., unpublished data). *lit-1(or131)* mutant embryos have a highly penetrant *mom*-like defect with 249/252 mutant embryos lacking gut and making extra pharyngeal tissue at the restrictive temperature (25 °C) as scored by Nomarski optics. RNAi was performed on W06F12.b using the Kohara complementary DNA clone yk338a10. We sequenced all mutant alleles as described⁵. Two-hybrid cloning of *tap-1* was performed as described²⁸. A full-length *mom-4* cDNA was fused to the LexA DNA-binding domain in pBTM116. We screened a GAL4 activation-domain cDNA library from mixed stage *C. elegans*. Interactions were detected in the yeast strain L40 on media lacking histidine.

Kinase assays. The mammalian expression vectors were transfected into 293 cells by using the calcium phosphate precipitation method. Proteins from the cell lysates were immunoprecipitated with respective antibodies. The immunoprecipitated proteins were analysed by immunoblotting and visualized by enhanced chemiluminescence (ECL, Amersham). Aliquots of immunoprecipitates were incubated with bacterially expressed His-MKK6 or GST-cJun in kinase buffer containing 10 mM HEPES (pH 7.4), 1 mM DTT, 5 mM MgCl₂ and 5 μ Ci of [γ -³²P]ATP at 25 °C for 2 min. Samples were analysed by SDS-PAGE and autoradiography.

Received 29 April; accepted 27 May 1999.

- Cadigan, K. M. & Nusse, R. Wnt signaling: a common theme in animal development. *Genes Dev.* **11**, 3286–3305 (1997).
- Lin, R., Thompson, S. & Priess, J. R. *pop-1* encodes an HMG box protein required for the specification of a mesoderm precursor in early *C. elegans* embryos. *Cell* **83**, 599–609 (1995).
- Lin, R., Hill, R. & Priess, J. R. POP-1 and anterior-posterior fate decision in *C. elegans* embryos. *Cell* **92**, 229–240 (1998).
- Rocheleau, C. E. et al. Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. *Cell* **90**, 707–716 (1997).
- Thorpe, C. J., Schlesinger, A., Carter, J. C. & Bowerman, B. Wnt signaling polarizes an early *C. elegans* blastomere to distinguish endoderm from mesoderm. *Cell* **90**, 695–705 (1997).
- Ishitani, T. et al. The TAK1–NLK–MAPK-related pathway antagonizes signalling between β -catenin and transcription factor-TCF. *Nature* **399**, 798–802 (1999).
- Goldstein, B. Induction of gut in *Caenorhabditis elegans* embryos. *Nature* **357**, 255–257 (1992).
- Goldstein, B. An analysis of the response to gut induction in the *C. elegans* embryo. *Development* **121**, 1221–1236 (1995).
- Brunner, E., Peter, O., Schweizer, L. & Basler, K. *pangolin* encodes a Lef-1 homolog that acts downstream of Armadillo to transduce the Wingless signal. *Nature* **385**, 829–833 (1997).
- Riese, J. et al. LEF-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic. *Cell* **88**, 777–787 (1997).
- van de Wetering, M. et al. Armadillo co-activates transcription driven by the product of the *Drosophila* segment polarity gene *TCF*. *Cell* **88**, 788–799 (1997).
- Bienz, M. TCF: transcriptional activator or repressor? *Curr. Opin. Cell Biol.* **10**, 366–372 (1998).
- Brannon, M., Gomperts, M., Sumoy, L., Moon, R. T. & Kimelman, D. A β -catenin/XTcf-3 complex binds to the siamois promoter to regulate dorsal axis specification in *Xenopus*. *Genes Dev.* **11**, 2359–2370 (1997).
- Cavallo, R. A. et al. *Drosophila* Tcf and Groucho interact to repress Wingless signalling activity. *Nature* **395**, 604–608 (1998).
- Roose, J. et al. The *Xenopus* Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature* **395**, 608–612 (1998).
- Kaletta, T., Schnabel, H. & Schnabel, R. Binary specification of the embryonic lineage in *Caenorhabditis elegans*. *Nature* **390**, 294–298 (1997).
- Yamauchi, K. et al. Identification of a member of the MAPKKK family as a potential mediator of TGF- β signal transduction. *Science* **270**, 2008–2011 (1995).
- Shibuya, H. et al. TAB1: an activator of the TAK1 MAPKKK in TGF- β signal transduction. *Science* **272**, 1179–1182 (1996).
- Fire, A., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
- Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extra-chromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).
- Brott, B. K., Pinsky, B. A. & Erikson, R. L. Nlk is a murine protein kinase related to Erk/MAP kinases and localized in the nucleus. *Proc. Natl Acad. Sci. USA* **95**, 963–968 (1998).
- Choi, K. W. & Benzer, S. Rotation of photoreceptor clusters in the developing *Drosophila* eye requires the nemo gene. *Cell* **78**, 125–136 (1994).
- Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
- Epstein, H. F. & Shakes, D. C. *Caenorhabditis elegans: Modern Biological Analysis of an Organism* (Academic, San Diego, 1995).
- Avery, L., Horvitz, H. R. Pharyngeal pumping continues after laser killing of the pharyngeal nervous system of *C. elegans*. *Neuron* **3**, 473–485 (1989).
- Bowerman, B., Eaton, B. A. & Priess, J. R. *skn-1*, a maternally expressed gene required to specify the fate of ventral blastomeres in the early *C. elegans* embryo. *Cell* **68**, 1061–1075 (1992).
- Priess, J. R. & Thomson, J. N. Cellular interactions in early *C. elegans* embryos. *Cell* **34**, 85–100 (1987).
- Fields, S. & Sternglanz, R. The two-hybrid system: an assay for protein–protein interactions. *Trends Genet.* **8**, 286–292 (1994).

Acknowledgements. We thank J. Willis for assistance with POP-1 antibody staining; D. Miller, R. Lin and J. Ahn for providing antibodies; the *C. elegans* Genetics center, funded by the NIH National Center for Research Resources, for providing strains; H. and R. Schnabel for the *t1534 lit-1* allele; and B. Draper and C. Doe for helpful comments on the manuscript. B.B. is supported by an NIH R01 grant and M.M. by an NIH Genetics training grant.

Correspondence and requests for materials should be addressed to B.B. (e-mail: bbowerman@molbio.uoregon.edu). Genbank accession numbers are: MOM-4, AF145377; TAP-1, AF145375; LIT-1, AF145376.

The TAK1–NLK–MAPK-related pathway antagonizes signalling between β -catenin and transcription factor TCF

Tohru Ishitani*, Jun Ninomiya-Tsuji*, Shin-ichi Nagai†, Michiru Nishita†, Marc Meneghini‡, Nick Barker§, Marian Waterman||, Bruce Bowerman‡, Hans Clevers§, Hiroshi Shibuya† & Kunihiro Matsumoto*

* Department of Molecular Biology, Graduate School of Science, Nagoya University, and CREST, Japan Science and Technology Corporation, Chikusa-ku, Nagoya 464-8602, Japan

† Division of Morphogenesis, Department of Developmental Biology, National Institute for Basic Biology, Okazaki 444-8585, Japan

‡ The Institute of Molecular Biology, The University of Oregon, 1370 Franklin Blvd., Eugene, Oregon 97403, USA

§ Department of Immunology, University Hospital, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands

|| Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, California 92697-4025, USA

The Wnt signalling pathway regulates many developmental processes through a complex of β -catenin and the T-cell factor/lymphoid enhancer factor (TCF/LEF) family of high-mobility-group transcription factors^{1–5}. Wnt stabilizes cytosolic β -catenin, which then binds to TCF and activates gene transcription. This signalling cascade is conserved in vertebrates, *Drosophila* and *Caenorhabditis elegans*. In *C. elegans*, the proteins MOM-4 and LIT-1 regulate Wnt signalling to polarize responding cells during embryogenesis⁶. MOM-4 and LIT-1 are homologous to TAK1 (a

kinase activated by transforming growth factor- β) mitogen-activated protein-kinase-kinase (MAP3K)⁷ and MAP kinase (MAPK)-related NEMO-like kinase (NLK)^{8,9}, respectively, in mammalian cells. These results raise the possibility that TAK1 and NLK are also involved in Wnt signalling in mammalian cells. Here we show that TAK1 activation stimulates NLK activity and downregulates transcriptional activation mediated by β -catenin and TCF. Injection of NLK suppresses the induction of axis duplication by microinjected β -catenin in *Xenopus* embryos. NLK phosphorylates TCF/LEF factors and inhibits the interaction of the β -catenin–TCF complex with DNA. Thus, the TAK1–NLK–MAPK-like pathway negatively regulates the Wnt signalling pathway.

To investigate whether TAK1 can influence Wnt signal transduction, we analysed the effects of TAK1 on the transcriptional activities of a β -catenin–TCF complex using two luciferase reporter constructs containing either multimeric TCF-binding sites (TOPFLASH) or mutated TCF-binding sites (FOPFLASH)^{10,11}. Regulation of β -catenin turnover requires the amino-terminal region of the protein. Deletion of this sequence (β -catenin Δ N) results in the accumulation of β -catenin and thus activates its signalling activity¹². Co-transfection of β -catenin Δ N and the reporter TOPFLASH into human embryonic kidney 293 cells resulted in about a 25-fold increase in luciferase activity relative to TOPFLASH alone (Fig. 1a). As expected, no activity was observed when the mutated reporter construct FOPFLASH was co-transfected. Overexpression of TAK1 together with the binding protein TAB1 activates TAK1 kinase activity¹³. Co-transfection of TAB1 and TAK1 together with β -catenin Δ N into 293 cells suppressed the β -catenin Δ N-induced transcriptional activation (Fig. 1a). Co-transfection of β -catenin Δ N, TAB1 and a kinase-inactive mutant of TAK1, TAK1(K63W) (ref. 7), did not block but slightly enhanced TCF-dependent transcription. Thus, TAK1-mediated inhibition of the TCF response depends on its kinase activity. Repression by TAB1–TAK1 was specific, as they had no effect on the basal activity of the mutant TCF reporter gene FOPFLASH. These results indicate that TAK1 negatively regulates the Wnt pathway.

Next, we examined the effect of NLK on Wnt signalling. In transient transfections using the TCF reporter-gene assay, overexpression of NLK repressed the activation of transcription by β -catenin Δ N in a dose-dependent manner (Fig. 1b). A kinase-inactive mutant of NLK, NLK(K155M) (ref. 9), partially blocked activation of the response to TCF mediated by β -catenin Δ N.

The Wnt pathway is crucial in the development of the *Xenopus* embryonic axis^{14,15}. We therefore analysed the *in vivo* function of NLK using the axis-formation assay in *Xenopus* embryos. We injected *in vitro*-synthesized NLK messenger RNA into the dorsal or ventral sides of four-cell-stage embryos. Injection of NLK mRNA into the dorsal marginal zone inhibited formation of the endogenous axis, resulting in ventralization of the embryos (Fig. 2a). To test whether NLK acts through the Wnt pathway, we examined the effects of NLK on axis duplication induced by β -catenin, *Siamois* or *Twin*. Injection of β -catenin into the vegetal ventral region of early cleavage-stage embryos leads to axis duplication¹⁶. *Siamois* and *Twin* are homeobox genes whose expression is specifically activated by β -catenin–TCF and which mediate the effects of the Wnt pathway on axis formation^{17–19}. Co-injection of NLK strongly suppressed β -catenin-induced axis duplication and, in most cases, resulted in normal development of the embryos (Fig. 2b, c). NLK could not block axis duplication induced by injection of *Siamois* or *Twin* mRNA (Fig. 2c). Thus, injection of NLK mRNA can block either normal or secondary dorsal-axis formation in *Xenopus* embryos, apparently by interfering with signalling through the Wnt pathway at a point downstream of β -catenin, and upstream of *Siamois* and *Twin*. These results indicate that NLK is a negative regulator of the Wnt pathway.

TAK1 is a member of the MAP3K family⁷, and NLK is related to

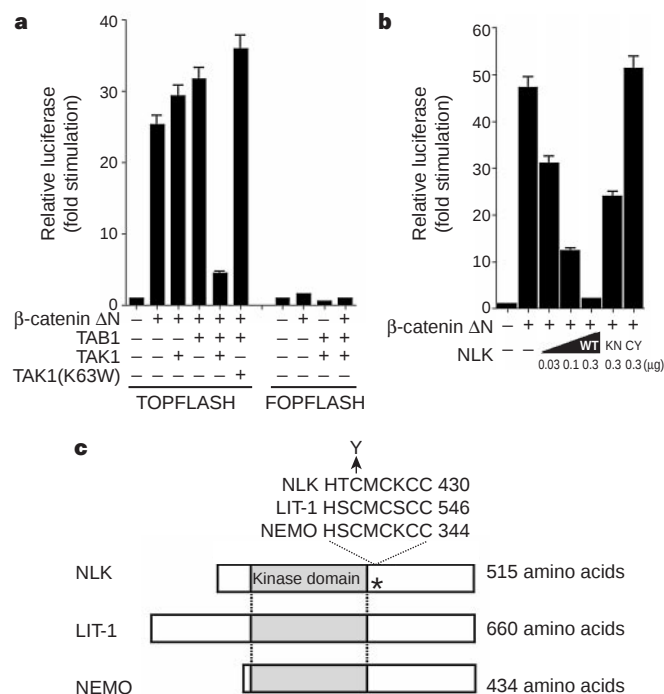


Figure 1 Effect of TAK1 and NLK on β -catenin–TCF-regulated transcription. **a**, **b**, 293 cells were transfected with luciferase reporter plasmid and expression vectors as indicated (**b**, NLK: WT, wild type; KN, NLK(K155M); CY, NLK(C425Y)). Relative luciferase activity was measured. **c**, Mutation site of NLK(C425Y). Diagrams of the structures of NLK, *C. elegans* LIT-1 and *Drosophila* NEMO are shown. The NLK(C425Y) mutation site is indicated by an asterisk. The amino-acid residues in the region of the mutation site are aligned above the diagram.

the downstream MAPK and is localized in the nucleus⁹, indicating that TAK1 may function upstream of NLK to negatively regulate the Wnt pathway. To test this possibility, we determined whether activation of TAK1 increases NLK kinase activity. We transiently expressed a Flag-epitope-tagged wild-type (Flag-NLK) or kinase-inactive mutant of NLK (Flag-NLK(K155M)) in 293 cells. The epitope-tagged proteins were immunoprecipitated with an antibody raised against Flag, and incubated with [γ -³²P]ATP (Fig. 3a). Wild-type NLK was autophosphorylated, and NLK(K155M) was not. When Flag-NLK was co-transfected with TAB1 and TAK1, TAK1 could stimulate NLK autophosphorylation activity, indicating that TAK1 acts upstream of NLK to activate its kinase activity.

We analysed the levels of β -catenin and its localization in 293 or COS7 cells co-transfected with β -catenin and NLK. There were no detectable differences in the intracellular distribution or amounts of β -catenin in these cells compared to those transfected with only β -catenin (data not shown). Thus, the inhibition of β -catenin–TCF signalling by NLK is not attributable to any effect on nuclear localization or the intracellular amounts of β -catenin. These results indicate that the TAK1–NLK pathway may affect Wnt signalling through TCF. To test this possibility, we examined whether NLK associated with TCF-4. We co-transfected Flag-NLK and TCF-4 into 293 cells and then immunoprecipitated the lysates. TCF-4 was detected in NLK immunoprecipitates (Fig. 4a, lanes 1–3).

To determine whether NLK phosphorylates TCF-4, we examined the electrophoretic mobility of TCF-4. Western blot analysis revealed that TCF-4 proteins migrated more slowly on SDS–PAGE in cells co-expressing wild-type NLK (Fig. 4b, lane 2). When TAK1 and TAB1 were co-expressed with NLK, the amounts of more slowly migrating TCF-4 were greatly increased (Fig. 4b, lanes 4, 6). These slowly migrating bands were eliminated by phosphatase treatment (Fig. 4b, lane 7), indicating that they may represent phosphorylated TCF-4. Furthermore, expression of the kinase-negative mutant NLK(K155M) inhibited modification of TCF-4 even in the presence of TAB1–TAK1 (Fig. 4b, lane 5). Thus, activation of TAK1 induces phosphorylation of TCF-4 through NLK.

Next we tested whether NLK could phosphorylate TCF-4, using *in vitro* kinase assays. Cells expressing Flag-NLK and TCF-4 were immunoprecipitated with anti-Flag antibody and the precipitates were incubated with [γ -³²P]ATP (Fig. 4c). In these assays, both NLK

and TCF-4 became phosphorylated *in vitro* in an NLK-dependent manner (lanes 1, 2). Activation of TAK1 by co-expression with TAB1 stimulated phosphorylation of TCF-4 (lane 3), consistent with the hypothesis that TAK1 acts upstream to activate NLK kinase activity.

To obtain further support for the interaction of NLK with TCF, we tested whether NLK interacts with and phosphorylates other members of the TCF/LEF family. We co-transfected Flag-NLK and a haemagglutinin (HA)-epitope-tagged XTCTF-3 (HA–XTCTF-3) into 293 cells and then immunoprecipitated the lysates (Fig. 4a, lanes 5–7). XTCTF-3 was detected in NLK immunoprecipitates and NLK induced modification of XTCTF-3. When HA–LEF-1 was co-expressed with Flag-NLK, western blot analysis revealed that LEF-1 proteins migrated more slowly on SDS–PAGE (Fig. 4b, lane 9). Co-expression of TAK1 and TAB1 with NLK enhanced the mobility shift of LEF-1 (lanes 11, 13). This slowly migrating band was eliminated

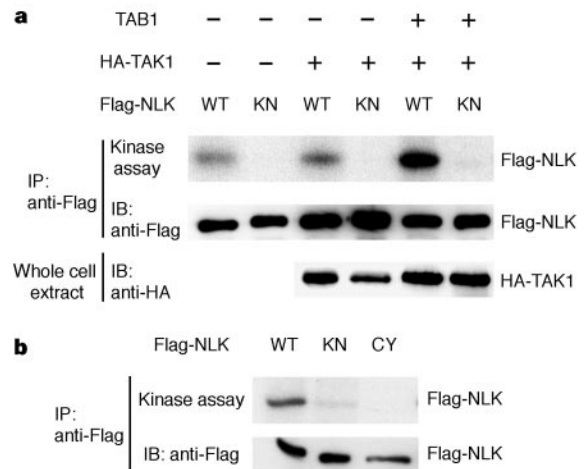


Figure 3 Kinase activity of NLK. **a**, **b**, 293 cells were transfected with expression vectors as indicated. Immunoprecipitated (IP) complexes with anti-Flag were incubated with [γ -³²P]ATP and analysed by autoradiography. The immunoprecipitates were also immunoblotted (IB) with anti-Flag. Expression of HA-TAK1 was monitored (**a**).

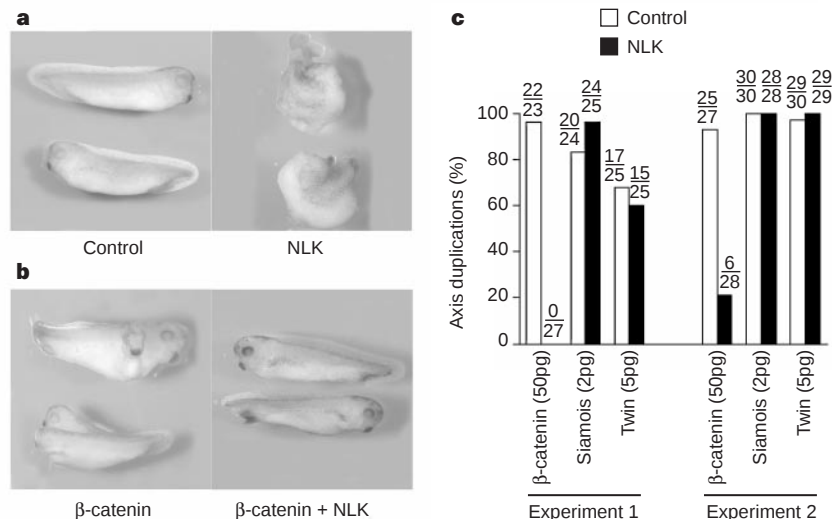


Figure 2 Effects of NLK on axis formation. **a**, NLK mRNA (250 pg) was injected into each of two dorsal blastomeres at the four-cell stage. Embryos were evaluated at the tadpole stage, and examples are shown. The average dorso-anterior index (DAI), a measure of the degree of dorsal and anterior mesodermal patterning, for each group was: control, average DAI = 5.0 (n = 31); NLK, average DAI = 3.1

(n = 30). **b**, **c**, mRNA encoding the indicated dorsalizing factor (β -catenin, Siamois or Twin) was injected into two ventral blastomeres at the four-cell stage with or without NLK mRNA (250 pg). Embryos were evaluated at the tadpole stage, and examples are shown (**b**). The fraction of embryos with duplicated axes is indicated above each bar (**c**).

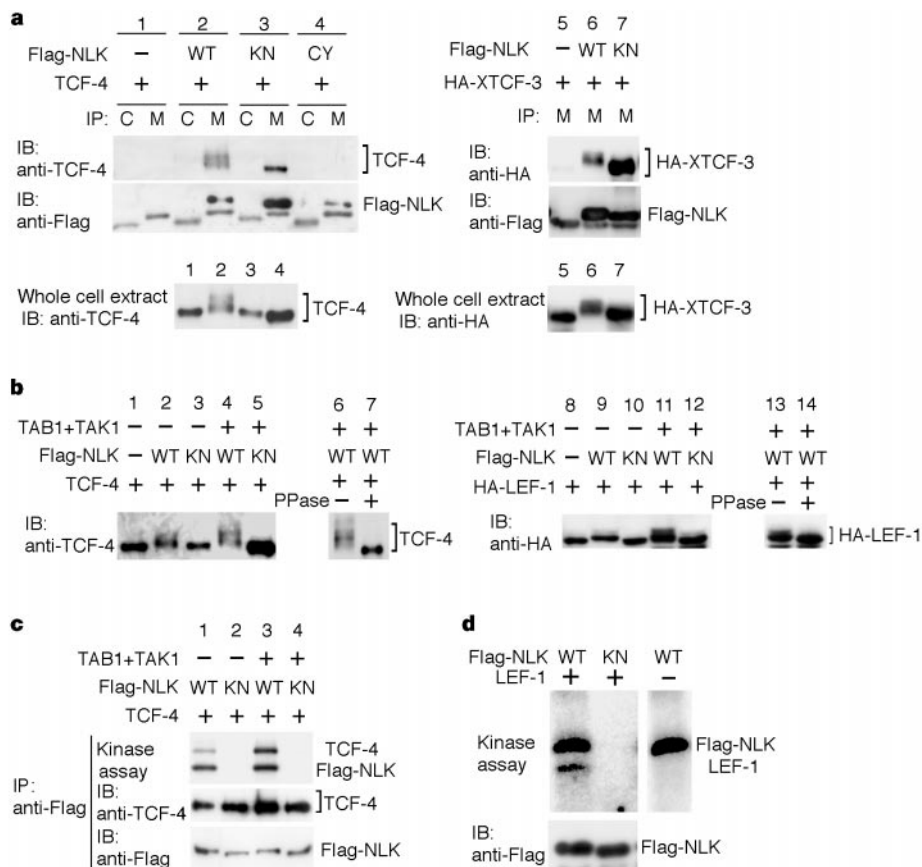


Figure 4 Interaction between TCF/LEF and the TAK1-NLK pathway. **a**, Association of NLK with TCF-4 and XTcf-3. 293 cells were transfected with expression vectors as indicated. Cell lysates were immunoprecipitated with anti-Flag (M) or IgG (C). Immunoprecipitates were immunoblotted with each antibody. Expression of TCF-4 or HA-XTCF-3 was monitored. **b**, Phosphorylation of TCF-4 and LEF-1. 293 cells were transfected with expression vectors as indicated. Cell lysates were incubated without or with protein phosphatase (PPase). Anti-TCF-4 or anti-HA

was used to detect TCF-4 or HA-LEF-1, respectively, in total cell lysates. **c**, **d**, Phosphorylation of TCF/LEF family proteins by NLK *in vitro*. 293 cells were transfected with expression vectors as indicated. Immunoprecipitated complexes with anti-Flag were incubated with [γ - 32 P]ATP (**c**) or with bacterially purified LEF-1 protein and [γ - 32 P]ATP (**d**). The immunocomplexes were immunoblotted with each antibody.

by phosphatase treatment (lane 14). Thus, NLK induces phosphorylation of LEF-1 *in vivo*. We also used bacterially expressed, highly purified LEF-1 as a substrate for *in vitro* kinase assays. Flag-NLK immunoprecipitated from 293 cells could phosphorylate LEF-1 (Fig. 4d), indicating that NLK may phosphorylate LEF-1 proteins directly. These findings indicate that NLK can associate with and phosphorylate TCF/LEF proteins.

A *C. elegans* *lit-1* loss-of-function mutant changes Cys 541 to a tyrosine within the conserved part of the LIT-1 C terminus (Fig. 1c)⁶. We introduced the analogous mutation into NLK to generate NLK(C425Y) and compared its biological activity to that of its wild-type counterpart. The NLK(C425Y) mutant could not block β -catenin-mediated transcriptional activation (Fig. 1b), and its kinase activity was dramatically decreased compared with that of wild-type NLK (Fig. 3b). Furthermore, co-immunoprecipitation analysis showed that the mutant NLK(C425Y) did not associate with TCF-4 (Fig. 4a, lane 4). These defects in NLK(C425Y) were not due to any significant protein instability, because the amounts of NLK(C425Y) expressed in 293 cells appeared to be roughly normal by immunoblot analysis. Thus, as in the *C. elegans* NLK homologue LIT-1, the Cys 425 residue of NLK is important for its function.

To investigate how NLK inhibits β -catenin–TCF-mediated transcriptional activation, we examined whether NLK interferes with β -catenin binding to TCF (Fig. 5a). We co-expressed β -catenin and TCF-4 in the presence or absence of Flag-NLK in 293 cells. Cell lysates were immunoprecipitated with anti- β -catenin antibody, and

co-precipitated TCF-4 was examined by immunoblot analysis. NLK did not affect the interaction between β -catenin and TCF-4. Flag-NLK also co-precipitated with β -catenin, indicating that NLK may form a complex with β -catenin and TCF-4.

We next asked whether NLK affects the DNA-binding properties of the β -catenin–TCF complex (Fig. 5b). In cell extracts from 293 cells co-transfected with β -catenin and TCF-4, we detected TCF-4 and β -catenin–TCF-4 by a gel-retardation-supershift assay using the TCF-binding motif as a probe. There was no binding to an oligonucleotide containing mutated TCF-binding sites. NLK over-expression led to a decrease in the amount of β -catenin–TCF-4–DNA complex, but had little effect on that of TCF-4–DNA. A catalytically inactive NLK(K155M) had no effect on complex binding, indicating that the effect of NLK depends on its kinase activity. These results indicate that NLK phosphorylates TCF and then interferes with the binding of β -catenin–TCF to the TCF target sites, thereby negatively regulating β -catenin–TCF-mediated events.

Regulation of TCF as both an activator and a repressor may reflect a balance between TCF with and without β -catenin^{1–5,20–22}. NLK may provide a negative control mechanism that discriminates between TCF bound alone as a repressor and TCF bound in association with β -catenin. In *C. elegans*, both LIT-1 and the β -catenin homologue WRM-1 are required to downregulate the TCF-like protein POP-1 (refs 6, 23). This raises the possibility that association between NLK and TCF/LEF *in vivo* requires β -catenin.

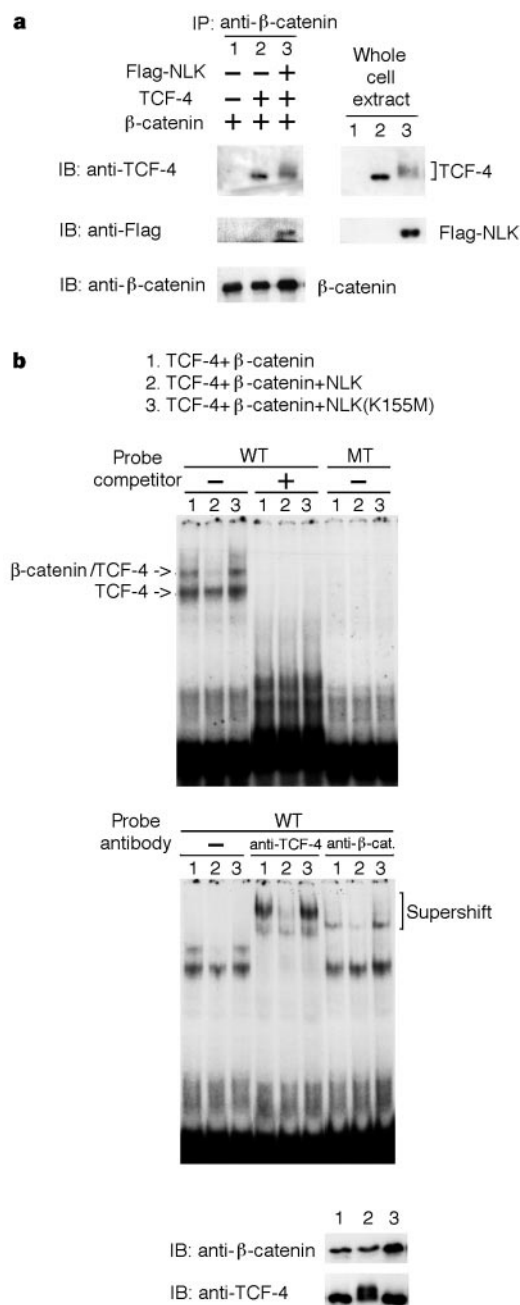


Figure 5 Effect of NLK on a β -catenin-TCF-4 complex. **a**, 293 cells were transfected with expression vectors as indicated. Cell lysates were immunoprecipitated with anti- β -catenin. The immunoprecipitates were immunoblotted with each antibody. Expression of TCF-4 and Flag-NLK was monitored. **b**, Gel-retardation-supershift assays were performed on cell extracts from 293 cells transfected with expression vectors as indicated. Samples were incubated with the optimal TCF retardation probe (WT) or the mutant (non-binding) probe (MT). Excess unlabelled TCF probe (competitor) was incubated in the reaction. Antibodies were added as indicated. Expression of β -catenin and TCF-4 was monitored.

Consistent with this, we detected no interaction of NLK with TCF-1, TCF-3, TCF-4 or LEF-1 in yeast two-hybrid assays, indicating that NLK may not interact with TCF/LEF directly. Furthermore, endogenous NLK interacts with both TCF-4 and β -catenin in the human colorectal tumour cell line SW480 (data not shown), which contains a stable β -catenin-TCF complex resulting from inactivation of the adenomatous polyposis coli tumour-suppressor gene^{24,25}. Thus, β -catenin may be required to make a stable complex between NLK and TCF/LEF.

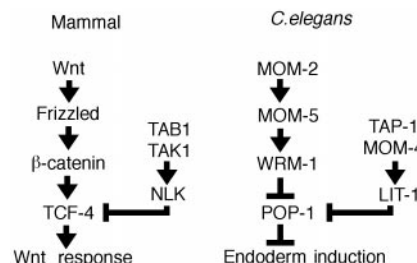


Figure 6 Model for interaction between the Wnt and TAK1/MOM-4-NLK/LIT-1 pathways. See text for details. *C. elegans* TAP-1 (TAB1-like protein) was identified as an MOM-4-binding protein and functioned as an activator of MOM-4 (ref. 6).

Our studies reveal conservation between mammals and *C. elegans* in the downregulation of HMG-domain proteins through MAPK-like pathways (Fig. 6). Studies in mammalian cells indicate that Wnt signalling stabilizes cytoplasmic β -catenin, which translocates to the nucleus and converts TCF-like repressors into transcriptional activators^{1-5,20-22}. In *C. elegans*, WRM-1 appears to block the repressor function of POP-1, rather than converting it into an activator^{23,26,27}. These different modes of action may be the basis for the observation that the *C. elegans* MOM-4-LIT-1 pathway positively regulates Wnt signalling, whereas the mammalian TAK1-NLK pathway negatively regulates it (Fig. 6). We propose a model in which TAK1/MOM-4 activates NLK/LIT-1, which in turn phosphorylates TCF/POP-1, leading to inhibition of the ability of TCF/POP-1 to activate or repress transcription. Thus, interactions between the Wnt signal-transduction pathway and a MAPK-related pathway are conserved between *C. elegans* and vertebrates. It will be interesting to determine how much of Wnt signalling during animal development is subject to modulation by MAP-kinase-related activities. □

Methods

Reporter gene assays. 293 cells (1.6×10^5 cells per well) were seeded into six-well (35 mm) plates. Cells were transfected by the calcium phosphate precipitate method at 24 h after seeding with the TOPFLASH or FOPFLASH reporter gene plasmid together with each expression vector, as indicated. The total amount of DNA (1.7 μ g) was kept constant by supplementation with empty vector DNAs. We determined luciferase activity with a Luciferase Assay System (Promega). We used the β -Gal vector (0.1 μ g) under the control of the β -actin promoter for normalizing transfection efficiencies.

Axis formation assays. The NLK cDNA was subcloned into the CS2+ vector²⁸. We synthesized capped mRNA from linearized vectors using a Message Machine kit (Ambion). The RNA was then injected into the marginal zones of early-stage embryos.

Immunoprecipitation. The immunoprecipitates and aliquots of total lysates were resolved on SDS-PAGE, and transferred to PVDF membranes (Hybond-P, Amersham). The membranes were immunoblotted with antibodies, and bound antibodies were visualized with horseradish-peroxidase-conjugated antibodies to mouse IgG using the Enhanced Chemiluminescence (ECL) Western Blotting System (Amersham).

In vitro kinase assays. We incubated aliquots of immunoprecipitates in 10 μ l kinase buffer containing 10 mM HEPES (pH 7.4), 1 mM DTT, 5 mM MgCl₂, and 5 μ Ci of [γ -³²P]ATP at 25 °C for 2 min. Samples were resolved on SDS-PAGE, and phosphorylated proteins were visualized by autoradiography.

Gel-retardation assays. As the optimal TCF probe, we used a double-stranded 56-nucleotide oligomer containing three potential TCF/LEF-binding sites derived from TOPFLASH. The mutant probe was derived from FOPFLASH. Binding reactions were done at room temperature for 10 min by incubating 6 μ l total lysates and 0.035 pmol labelled oligonucleotides in 15 μ l binding buffer (10 mM Tris pH 7.5, 50 mM NaCl, 1 mM MgCl₂, 0.5 mM EDTA, 0.5 mM DTT, 4% glycerol, 0.75 μ g poly(dI-dC), 1.5 μ g salmon sperm DNA and 1.5 μ g BSA). Competition analyses were performed with excess (1.0 pmol) unlabelled probe.

Received 29 April; accepted 27 May 1999.

- Behrens, J. *et al.* Functional interaction of β -catenin with the transcription factor Lef-1. *Nature* **382**, 638–642 (1996).
- Molenaar, M. *et al.* XTcf-3 transcription factor mediates β -catenin-induced axis formation in *Xenopus* embryos. *Cell* **86**, 391–399 (1996).
- Brunner, E., Peter, O., Schweizer, L. & Basler, K. *pangolin* encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in *Drosophila*. *Nature* **385**, 829–833 (1997).
- Riese, J. *et al.* Lef-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic. *Cell* **88**, 777–787 (1997).
- van de Wetering, M. *et al.* Armadillo coactivates transcription driven by the product of the *Drosophila* segment polarity gene DTCF. *Cell* **88**, 789–799 (1997).
- Meneghini, M. D. *et al.* MAP kinase and Wnt pathways converge to downregulate an HMG-domain repressor in *Caenorhabditis elegans*. *Nature* **399**, 793–797 (1999).
- Yamaguchi, K. *et al.* Identification of a member of the MAPKKK family as a potential mediator of TGF- β signal transduction. *Science* **270**, 2008–2011 (1995).
- Choi, K. W. & Benzer, S. Rotation of photoreceptor clusters in the developing *Drosophila* eye requires the *nemo* gene. *Cell* **78**, 125–136 (1994).
- Brott, B. K., Pinsky, B. A. & Erikson, R. L. Nlk is a murine protein kinase related to Erk/MAP kinases and localized in the nucleus. *Proc. Natl Acad. Sci. USA* **95**, 963–968 (1998).
- van de Wetering, M., Oosterwegel, M., Dooijes, D. & Clevers, H. Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *EMBO J.* **10**, 123–132 (1991).
- van de Wetering, M., Castrop, J., Korinek, V. & Clevers, H. Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. *Mol. Cell Biol.* **16**, 745–752 (1996).
- Aberle, H. *et al.* β -Catenin is a target for the ubiquitin-proteasome pathway. *EMBO J.* **16**, 3797–3804 (1997).
- Shibuya, H. *et al.* TAB1: an activator of the TAK1 MAPKKK in TGF- β signal transduction. *Science* **272**, 1179–1182 (1996).
- Miller, J. R. & Moon, R. T. Signal transduction through β -catenin and specification of cell fate during embryogenesis. *Genes Dev.* **10**, 2527–2539 (1996).
- Cardigan, K. M. & Nusse, R. Wnt signalling: a common theme in animal development. *Genes Dev.* **11**, 3286–3305 (1997).
- Funayama, N., Fagotto, E., McCrea, P. & Gumbiner, B. M. Embryonic axis induction by the armadillo repeat domain of β -catenin: evidence for intracellular signaling. *J. Cell Biol.* **128**, 959–968 (1995).
- Lemaire, P., Garrett, N. & Gurdon, J. B. Expression cloning of *siamois*, a *Xenopus* homeobox gene expressed in dorsal vegetal cells of blastulae and able to induce a complete secondary axis. *Cell* **81**, 85–94 (1995).
- Brannon, M. *et al.* A β -catenin/XTcf-3 complex binds to the *siamois* promoter to regulate dorsal axis specification in *Xenopus*. *Genes Dev.* **11**, 2359–2370 (1997).
- Laurent, M. N. *et al.* The *Xenopus* homeobox gene *twin* mediates Wnt induction of gooseoid in establishment of Spemann's organizer. *Development* **124**, 4905–4916 (1997).
- Waltzer, L. & Bienz, M. *Drosophila* CBP represses the transcription factor TCF to antagonize Wingless signalling. *Nature* **395**, 521–525 (1998).
- Cavallo, R. A. *et al.* A *Drosophila* Tcf and Groucho interact to repress Wingless signalling activity. *Nature* **395**, 604–608 (1998).
- Roose, J. *et al.* The *Xenopus* Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature* **395**, 608–612 (1998).
- Rocheleau, C. E. *et al.* Wnt signaling and an APC related gene specify endoderm in early *C. elegans* embryos. *Cell* **90**, 707–717 (1997).
- Korinek, V. *et al.* Constitutive transcriptional activation by a β -catenin-Tcf complex in APC^{-/-} colon carcinoma. *Science* **275**, 1784–1787 (1997).
- Morin, P. J. *et al.* Activation of β -catenin-Tcf signaling in colon cancer by mutations in b-catenin or APC. *Science* **275**, 1787–1790 (1997).
- Thorpe, C. J., Schlesinger, A., Carter, J. C. & Bowerman, B. Wnt signaling polarizes in early *C. elegans* blastomeres to distinguish endoderm from mesoderm. *Cell* **90**, 695–705 (1997).
- Lin, R., Hill, R. J. & Priess, J. R. POP-1 and anterior–posterior fate decisions in *C. elegans* embryos. *Cell* **92**, 229–239 (1998).
- Rupp, R. A., Snider, H. & Weintraub, H. *Xenopus* embryos regulate the nuclear localization of XMyoD. *Genes Dev.* **8**, 1311–1323 (1994).

Acknowledgements. We thank T. Akiyama, B. Brott, and A. Kikuchi for materials; E. Nishida for helpful discussions; and M. Lamphier and R. Ruggieri for critical reading of the manuscript. Supported by special grants for CREST and Advanced Research on Cancer from the Ministry of Education, Culture and Science of Japan, and HFSP (K.M.).

Correspondence and requests for materials should be addressed to K.M. (e-mail: g44177@nuccl.cc.nagoya-u.ac.jp).

TRA-1 regulates the cellular distribution of the *tra-2* mRNA in *C. elegans*

L. E. Graves, S. Segal & E. B. Goodwin

Department of Cell and Molecular Biology and the Lurie Cancer Center, Northwestern University Medical School, Chicago, Illinois 60611, USA

The GLI protein family is involved in several key developmental processes in both vertebrates and invertebrates. The *Drosophila* GLI protein, Cubitus interruptus (Ci), regulates segment polarity and wing and leg development. In vertebrates, the GLI proteins control neural, lung, bone and gut development¹. In the nematode

Caenorhabditis elegans, the GLI family member TRA-1 is necessary for normal sexual development^{2,3}. GLI, Ci and TRA-1 each contain five zinc-finger domains and bind the identical DNA sequence. Previous analyses are consistent with these proteins being transcription factors^{4,5}. Here we show that TRA-1 can act post-transcriptionally to govern gene activity. Our results indicate that the binding of TRA-1 to the 3' untranslated region of *tra-2* regulates the export of *tra-2* messenger RNA from the nucleus. The fact that TRA-1 is part of a conserved family of proteins raises the possibility that GLI family members are both transcriptional and post-transcriptional regulators of gene expression.

C. elegans has two sexes—hermaphrodite and male. Hermaphrodites are essentially female animals that produce both sperm and oocytes. *tra-1* (*transformer-1*), the terminal regulator in the *C. elegans* sex-determination pathway, has an essential and primary role in female development (loss of *tra-1* activity transforms hermaphrodites into males) and a lesser role in male development^{2,3}. *tra-1* produces two proteins: TRA-1A and TRA-13 (ref. 6); for simplicity, we collectively refer to both proteins as TRA-1. *tra-2*, which acts upstream of *tra-1* in the sex-determination pathway, encodes a large transmembrane protein and is also required for proper female development^{7,8}. Northern analysis demonstrates that TRA-1 positively feeds back onto *tra-2* (ref. 9), as loss of *tra-1* activity decreases *tra-2* mRNA levels. This positive feedback loop might ensure that cells are fully committed to the female fate.

To investigate how *tra-1* promotes *tra-2* activity, we used RNA *in situ* analysis to compare the subcellular distribution of *tra-2* mRNA between wild-type and *tra-1*(null) animals. We examined the *tra-2* mRNA in the intestines of adult animals, as the sexual phenotype of this tissue is sensitive to changes in both *tra-1* and *tra-2* activity. We observe a marked difference in the distribution of

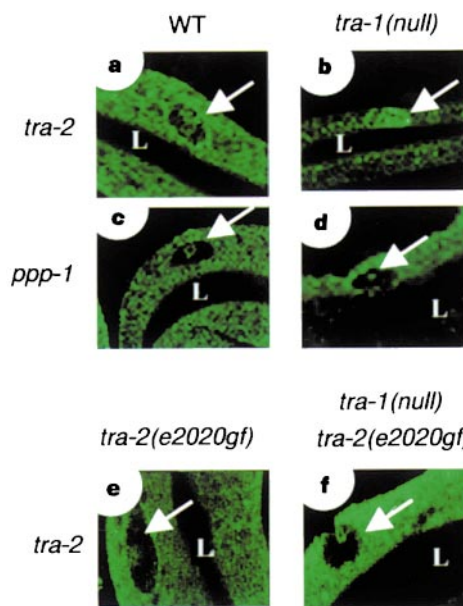


Figure 1 TRA-1 regulates the levels and localization of *tra-2* mRNA. RNA *in situ* hybridization of intestines of wild-type (wt) hermaphrodites and *tra-1*(*e1834null*) animals. *tra-1*(*e1834null*) is a molecular null². Wild-type and *tra-1*(null) animals were fixed and treated adjacent to one another on a common slide. **a**, In wild-type hermaphrodites, *tra-2* mRNA appears cytoplasmic. **b**, In *tra-1*(null) males, *tra-2* mRNA is predominantly nuclear with a dramatic decrease in cytoplasmic levels. The nucleolar signals are insignificant as controls show this pattern. In wild-type males, *tra-2* nuclear mRNA is slightly increased and the cytoplasmic mRNA levels are reduced, consistent with lower TRA-1 activity (data not shown). For both **c**, wild-type, and **d**, *tra-1*(null) animals, *ppp-1* mRNA is cytoplasmic. *tra-2*(*e2020gf*) mRNA is cytoplasmic in **e**, *tra-2*(*e2020gf*) and **f**, *tra-2*(*e2020gf*); *tra-1*(*e1834null*) animals. Arrows indicate nuclei. L, lumen.

Table 1 TRA-1 is required for the activity of β -gal reporter transgenes

Reporter transgene*	Animals with intestinal β -gal activity (%)†	
	Wild-type	<i>tra-1(e1834null)</i>
<i>lacZ::tra-2(+)</i> 3'UTR	29 (n = 69)	11 (n = 62)
<i>lacZ::tra-2(Δ32)</i> 3'UTR	74 (n = 71)	46 (n = 71)
<i>lacZ::tra-2(Δ60)</i> 3'UTR	70 (n = 56)	34 (n = 53)
<i>lacZ::tra-2(Δ108)</i> 3'UTR	56 (n = 66)	66 (n = 76)

Wild-type adult animals were *tra-1(e1834)/eT1* or *tra-1(e1834)/qC1* heterozygous XX hermaphrodites. *tra-1(e1834null)* animals were the homozygous siblings of the XX hermaphrodites. We obtained similar values for wild-type XO males. The allele *tra-1(e1834null)* was chosen as it is a large deletion of the gene and does not make RNA (our unpublished results), and hence is likely to be a molecular null.

* Four different reporter transgenes were used. See text for description of transgenes. In all experiments, adult transgenic worms were heat-shocked for 2 h at 33 °C and subsequently allowed to recover for 2 h at 20 °C before staining.

† Transgenic animals were scored as positive if blue precipitate was detectable in intestinal cells at 630 $\times$ magnification. Numbers represent the per cent of transgenic animals with blue precipitate in intestinal cells, and are the averages of at least four experiments.

tra-2 mRNA between the wild-type and *tra-1(null)* animals. In the wild-type samples the *tra-2* mRNA is predominantly cytoplasmic (Fig. 1a), whereas in the *tra-1(null)* animals there is a relative reduction in cytoplasmic *tra-2* mRNA and an increase in the amount of *tra-2* mRNA in the nucleus (Fig. 1b, arrow). Immunocytochemistry using polyclonal antibodies against the *tra-2* protein reveals that *tra-2* protein levels are reduced in the *tra-1(null)* animals compared with the wild type, indicating that the altered mRNA distribution affects *tra-2* translation (our unpublished data).

Because *tra-1* is predicted to encode a transcription factor⁶, we investigated whether the above findings could result from changes in *tra-2* transcription. We took advantage of the fact that *tra-2* is the second gene in a bi-cistron and shares a promoter with *ppp-1* (ref. 10). If the changes in the pattern of *tra-2* mRNA result from alterations in transcription, the *ppp-1* mRNA distribution should likewise change in *tra-1(null)* animals. However, *ppp-1* mRNA is cytoplasmic and is present in similar amounts in both wild-type and *tra-1(null)* animals (Fig. 1c, d). This indicates that *tra-1* primarily affects the *tra-2* mRNA distribution at the post-transcriptional level.

Previous work has shown that *tra-2* is translationally regulated by two elements, called TGEs (*tra-2* GLI elements) in its 3' untranslated region (3'UTR)¹¹. We therefore used a series of reporter transgenes to test whether *tra-1* acts through the *tra-2* 3'UTR to regulate the *tra-2* mRNA. These transgenes all encode β -galactosidase (β -gal) with transcription driven by the inducible *C. elegans* heat-shock promoter (*hsp16.41*), and contain either the wild-type *tra-2* 3'UTR or a *tra-2* 3'UTR lacking one or both TGEs (*lacZ::tra-2(+)*3'UTR, *lacZ::tra-2(Δ 32)*3'UTR, or *lacZ::tra-2(Δ 60)*3'UTR, respectively).

Comparing the activity of these reporter transgenes between wild-type and *tra-1(null)* animals, we found that for all three transgenes the number of animals with intestinal β -gal staining

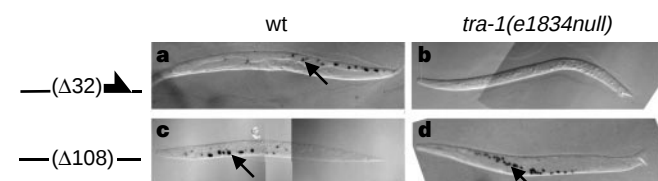


Figure 2 TRA-1 regulates reporter transgene activity through the *tra-2* 3'UTR. Lateral views of wild-type (a) and *tra-1(e1834null)* (b) animals carrying *lacZ::tra-2(Δ 32)*3'UTR (anterior is left). β -gal is detected in intestinal cells (arrow) in the wild-type hermaphrodite (a) but not in the *tra-1(null)* animal (b). Lateral views of wild-type (c) and *tra-1(null)* (d) animals carrying *lacZ::tra-2(Δ 108)*3'UTR (anterior is left). β -gal is detected in intestinal cells (arrow) of both wild-type and *tra-1(null)* animals. Left, schematics of the 3'UTRs carried by each transgene.

decreased in the absence of TRA-1 (Table 1 and Fig. 2: compare a and b). Only 11, 46 and 34% of *lacZ::tra-2(+)*3'UTR, *lacZ::tra-2(Δ 32)*3'UTR, and *lacZ::tra-2(Δ 60)*3'UTR, respectively, have intestinal β -gal activity in a *tra-1(null)* background, compared with 29, 74 and 70% for the same transgenes in a wild-type background (Table 1).

As observed with the endogenous *tra-2* mRNA, TRA-1 alters the localization of these transgene mRNAs. In wild-type hermaphrodites, *lacZ::tra-2(+)*3'UTR, *lacZ::tra-2(Δ 32)*3'UTR, and *lacZ::tra-2(Δ 60)*3'UTR mRNAs are predominantly cytoplasmic (Fig. 3a, and data not shown); however, in a *tra-1(null)* animal, the transgenic mRNA accumulates in the nucleus (Fig. 3b, arrow). These results are consistent with *tra-1* working through the *tra-2* 3'UTR to control the localization of the mRNA. In contrast to these findings, a fourth construct, lacking both TGEs and some flanking sequence (*lacZ::tra-2(Δ 108)*3'UTR), was insensitive to *tra-1* activity. The β -gal expression of *lacZ::tra-2(Δ 108)*3'UTR does not decrease in a *tra-1(null)* background (Fig. 2c, d, and Table 1), and RNA *in situ* analysis indicates that *lacZ::tra-2(Δ 108)*3'UTR mRNA is cytoplasmic regardless of *tra-1* activity (Fig. 3c, d, arrow). These results demonstrate a correlation between loss of β -gal expression and nuclear retention, and indicate that an element removed by the 108-nucleotide deletion is necessary for nuclear accumulation of mRNA. We therefore investigated whether the genetic *tra-2(e2020gf)* mutation, which deletes the same 108 nucleotides¹¹, also results in cytoplasmic mRNA regardless of *tra-1* activity. As with *lacZ::tra-2(Δ 108)*3'UTR, we found that *tra-2(e2020gf)* mRNA is cytoplasmic in hermaphrodite and *tra-1(null)* animals (Fig. 1e, f). Therefore, the 108-nucleotide deletion disrupts an element necessary for nuclear accumulation of the mRNA.

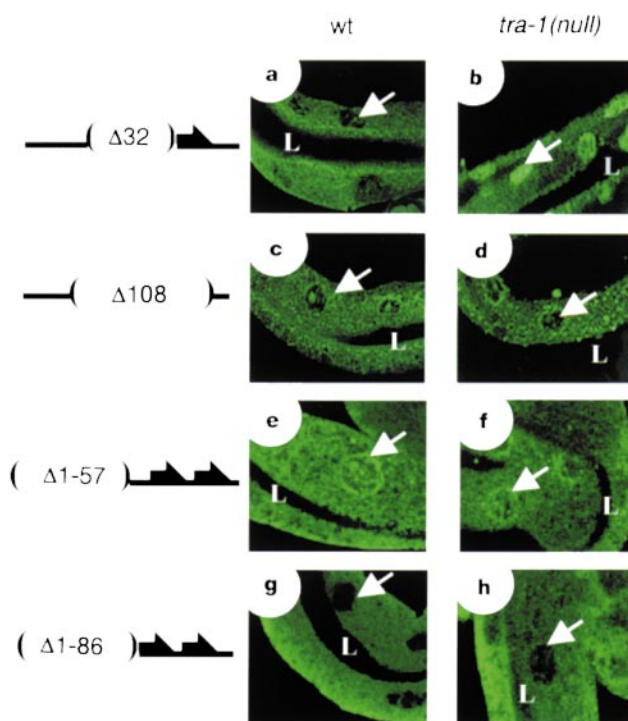


Figure 3 TRA-1 controls localization of reporter transgene mRNA. a, b, In wild-type animals (a), *lacZ::tra-2(Δ 32)*3'UTR mRNA is cytoplasmic, but in *tra-1(e1834null)* animals (b) it is nuclear. Nucleolar signals are insignificant, as they are present in controls. c, d, *lacZ::tra-2(Δ 108)*3'UTR RNA is cytoplasmic in both wild-type (c) and *tra-1(null)* (d) animals. e, f, *lacZ::tra-2(Δ 1-57)*3'UTR mRNA is predominately nuclear in both wild-type (e) and *tra-1(null)* (f) animals. g, h, *lacZ::tra-2(Δ 1-86)*3'UTR mRNA is cytoplasmic in both wild-type (g) and *tra-1(null)* (h) males. Left, schematics of the 3'UTRs carried by the transgenes. Arrows indicate nuclei.

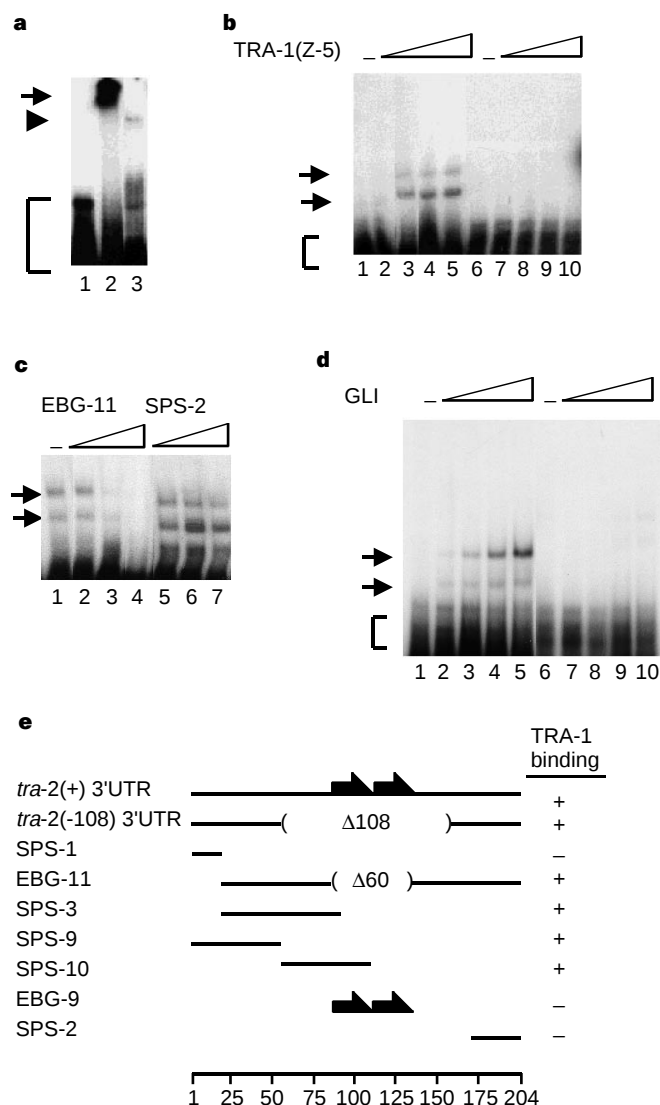


Figure 4 TRA-1 binds an element in the *tra-2* 3'UTR. **a**, Radioactively labelled *tra-2(+)*3'UTR (1 fmol) was incubated alone (lane 1), with 0.5 μ g of TRA-1(Z-5) (lane 2), or with 0.5 μ g of TRA-1(Z-2) (lane 3). Arrow and arrowhead indicate TRA-1(Z-5)-RNA and TRA-1(Z-2)-RNA complexes, respectively. Brackets indicate free probe. **b**, EBG-11 (1 fmol) (lanes 1–5) or SPS-2 (1 fmol) (lanes 6–10) incubated with 0 (lanes 1, 6), 0.15 (lanes 2, 7), 0.3 (lanes 3, 8), 0.6 (lanes 4, 9) or 0.9 μ g (lanes 5, 10) of TRA-1(Z-5). **c**, Labelled EBG-11 (1 fmol) was incubated with 0.3 μ g of TRA-1(Z-5) and either no competitor (lane 1) or 10, 50 or 100-fold molar excess of either cold EBG-11 (lanes 2–4) or cold SPS-2 (lanes 5–7). **d**, EBG-11 (1 fmol) (lanes 1–5) or SPS-2 (1 fmol) (lanes 6–10) incubated with 0, 0.1, 0.2, 0.5 and 1.0 μ g of truncated human GLI (arrow indicates complex). **e**, Summary of binding studies. Labelled RNAs (1 fmol each) were incubated with 0 to 1.0 μ g of TRA-1(Z-5). Names of RNAs are given on the left, schematics of RNAs in the middle; + and – indicate binding or no binding, respectively, at 0.3 μ g protein. The scale indicates *tra-2* 3'UTR nucleotide numbers. See Methods for probe sequences.

The above analyses indicate that *tra-1* works through the *tra-2* 3'UTR to regulate the subcellular localization of *tra-2* mRNA. The presence or absence of TGEs did not affect nuclear accumulation of transgenic mRNAs, and hence they are not essential for control by *tra-1*. The transgenic constructs lack introns, indicating that the *tra-1* effect is not likely to be brought about through splicing. Nuclear accumulation of the mRNA in the absence of TRA-1 is, however, dependent on an element in the *tra-2* 3'UTR, which we call a nuclear-localization element, that is disrupted by the 108-nucleotide deletion.

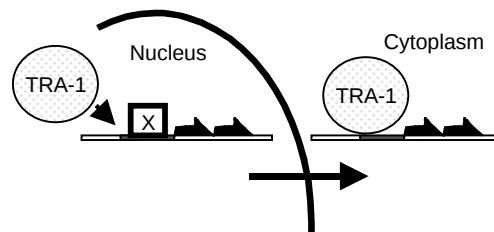


Figure 5 Model for the regulation of *tra-2* mRNA localization by TRA-1. Factor X binds the nuclear localization element in *tra-2* mRNA to retain it in the nucleus. TRA-1 binding to the *tra-2* 3'UTR releases the mRNA, allowing its export from the nucleus. TRA-1 may remain associated with the mRNA and stabilize it and/or alter its translation in the cytoplasm. Half-arrows indicate TGEs.

To determine whether TRA-1 regulates mRNA localization by directly binding the *tra-2* 3'UTR, we tested the ability of TRA-1 to bind the *tra-2* 3'UTR in RNA gel-shift experiments. Two constructs of TRA-1 were used for this analysis: an amino-terminal fragment that contained the five zinc-fingers (TRA-1(Z-5)) and an N-terminal fragment that has only the first two zinc-fingers (TRA-1(Z-2)). We found that both proteins bound the *tra-2(+)*3'UTR (Fig. 4a, lanes 2 and 3).

To identify the region of the *tra-2* 3'UTR required for complex formation, we assayed a series of deletions in the 3'UTR for their ability to bind TRA-1(Z-5). TRA-1(Z-5) specifically bound RNAs containing a small region of the *tra-2* 3'UTR 5' to the TGEs (called EBG-11 and SPS-3; Fig. 4b, lanes 1–5, and c). However, the protein did not bind the TGEs themselves (EBG-9; Fig. 4e), consistent with these elements not being required for TRA-1 regulation. Note that the TRA-1-binding site overlaps sequences removed by the 108-nucleotide deletion and may therefore partially coincide with the nuclear-localization element. It is of interest that the 3'UTR TRA-1-binding site has no apparent similarity to the TRA-1 DNA-binding site¹². Finally, as TRA-1 is a member of the GLI family, we tested whether human GLI (glioblastoma oncoprotein) bound the same sequences as TRA-1. A GLI fusion protein bound to an RNA containing the 5' region of the *tra-2* 3'UTR (Fig. 4d, lanes 2–5) but not to an RNA lacking this region (Fig. 4d, lanes 7–10). This indicates that regulation of RNA may be a conserved property of this protein family.

To examine whether there is a correlation between an intact TRA-1-binding site and the ability of the mRNA to be localized correctly, we constructed a transgene, called *lacZ::tra-2(Δ 1–57)3'UTR*, that carried a *tra-2* 3'UTR in which all sequences 5' to the 108-nucleotide breakpoint were deleted. We reasoned that removing part of the TRA-1-binding site, while maintaining the localization element, might impair TRA-1's ability to regulate mRNA localization, resulting in accumulation of transgenic mRNA in the nucleus of wild-type animals. As predicted, *lacZ::tra-2(Δ 1–57)3'UTR* mRNA accumulates in the nucleus of hermaphrodites (Fig. 3e, f), consistent with TRA-1 acting through the 3'UTR to control *tra-2* mRNA localization. This accumulation is not as dramatic as with *lacZ::tra-2(Δ 32)3'UTR*, suggesting that the deletion may have partially disrupted the nuclear-localization element. Removing an additional 3' 29 nucleotides (*lacZ::tra-2(Δ 1–86)3'UTR*) abolishes this nuclear accumulation (Fig. 3g, h), consistent with nucleotides 57–86 containing at least part of a nuclear-localization element.

In summary, we find that TRA-1 acts through the *tra-2* 3'UTR to regulate the subcellular localization of the *tra-2* mRNA. As the simplest model to explain our data, we propose that TRA-1 binding to the *tra-2* 3'UTR promotes transport of the *tra-2* mRNA out of the nucleus (Fig. 5). This process may have several roles in controlling sexual development. The export of *tra-2* mRNA may help to ensure that cells adopt the female fate. In addition, TRA-1 may enter the

cytoplasm with the *tra-2* mRNA and so titrate TRA-1's nuclear activity as a transcription factor. These functions may ensure proper sexual development by maintaining optimal TRA-2 and TRA-1 activity. A cytoplasmic role for TRA-1 is supported by immunocytochemistry experiments that find TRA-1 in the nucleus and cytoplasm (Y. Yu, L.E.G. and E.B.G., unpublished). The *tra-2* 3'UTR contains a nuclear-localization element that causes accumulation of the *tra-2* mRNA in the nucleus. Perhaps a nuclear factor binds this element, and TRA-1 competes with this factor for *tra-2* RNA binding. TRA-1 may export *tra-2* mRNA from the nucleus by a similar mechanism to that of the viral Rev protein, as TRA-1 contains putative REV-like NES sequences (E.B.G., unpublished results). The 3'UTR sequences required for regulation of mRNA localization are novel, as they have no similarity to previously identified sequences that control either nuclear retention or nuclear export of viral RNAs¹³. Furthermore, to our knowledge no cellular mRNA 3'UTR sequences have been described that control mRNA nuclear export. □

Methods

General procedures and strains. Strains were maintained as described¹⁴, and were raised at 20 °C. For all experiments, wild-type animals are *tra-1(e1834)/qC1* or *tra-1(e1834)/eT1* heterozygous XX hermaphrodites, and *tra-1(null)* animals are their homozygous XX siblings. The mutations and rearrangements used were as follows. Linkage group: LGII: *tra-2(e202gf)*; LGIII: *tra-1(e1834null)*. *qC1[dpy-19(e1259)glp-1(q339)]* and *eT1(unc-36)* suppress recombination over much of chromosomes III, or III and V, respectively.

RNA in situ analysis. RNA in situ hybridization experiments were performed on extruded intestines of adult animals as described¹⁵. To allow for comparison, wild-type and *tra-1(null)* animals were fixed and treated on a common slide, and photographed under identical conditions. The *tra-2* gene produces three mRNA transcripts, of 4.7, 1.9 and 1.8 kb¹⁶. The 1.9-kb and 1.8-kb transcripts are co-linear with the 3' end of the 4.7-kb mRNA¹⁷. To ensure all transcripts were examined, two different digoxigenin-labelled DNA probes were synthesized. All probes were made by primer extension. Probe A was specific to the 4.7-kb transcript and probe B recognized all three *tra-2* transcripts. Probe A was made by cutting pBG403 with *Bam*HI and the probe synthesized using the pGex 3' sequencing primer (Pharmacia). For the probe A, sense control, pBG403 was cut with *Eco*RI and the probe was synthesized with the pGEX 5' sequencing primer (Pharmacia). pBG403 contains a 774-bp fragment of the *tra-2* complementary DNA (nucleotides 121–895 as numbered in GenBank). Probe B was produced by cutting pJK111 with *Stu*I and the probe was synthesized with the T7 primer (Promega). For the probe B sense control, pJK111 was cut with *Sma*I and the probe was synthesized using the Sp6 primer (Promega). For *ppp-1*, 1.2 kb of the *ppp-1* cDNA was amplified from total *C. elegans* cDNA using primers LG-5 and LG-6 that contained *Kpn*I and *Sac*I restriction sites, respectively. The PCR product was cut with *Kpn*I and *Sac*I and cloned into Bluescript KSII cut with the same enzymes, making pBG404. For the *ppp-1* probe, pBG404 was cut with *Pvu*II and the probe was synthesized using the T7 primer (Promega). For the *ppp-1* sense control, pBG404 was linearized with *Pvu*II and synthesized using the T3 primer (Promega). To detect β -galactosidase reporter transgene mRNAs, pD1 (ref. 18), which contains an 800-bp fragment of the *lacZ* gene, was cut with *Hind*III and the probe synthesized using T7 primer (Promega). For the sense control probe, pD1 was cut with *Spe*I and synthesized with the T3 primer (Promega). All probes were digoxigenin-labelled and detected using a fluorescein-conjugated anti-digoxigenin antibody (Boehringer Mannheim).

Transgenic assays. The transgenes *lacZ::tra-2(+)*3'UTR, *lacZ::tra-2(Δ 60)*3'UTR, and *lacZ::tra-2(Δ 108)*3'UTR were made as described¹⁹. *lacZ::tra-2(Δ 1–57)*3'UTR and *lacZ::tra-2(Δ 1–86)*3'UTR were made by PCR, amplifying fragments from the plasmid pBG2 that contains the *tra-2(+)*3'UTR using the primers EJ-30 and either SPS-7 or SPS-5. EJ-30 contains an *Ap*I site, and SPS-7 and SPS-5 contain *Spe*I sites. The resulting PCR products were cut with *Spe*I and *Ap*I, and cloned into pPC16.41 (gift from P. Candido)¹⁹ cut with the same enzymes. All transgenic lines were constructed as described²⁰ with the injection solution containing either 25 ng μ l⁻¹ of test plasmid and 100 ng μ l⁻¹ of the plasmid pRF4 (which carries the dominant *rol-6* marker) or 15 ng μ l⁻¹ of

test plasmid and 150 ng μ l⁻¹ of pRF4. β -Galactosidase expression was assayed as described²¹.

RNA gel-shift assay. RNA gel-shift binding reactions were done as described¹¹. [α -³²P]-labelled and unlabelled RNA probes containing different 3'UTRs were produced by standard methods. Different full-length and mutant 3'UTRs were produced as described¹⁸. His-tagged TRA-1(Z-5) and TRA-1(Z-2) were produced using the His Binding Buffer Kit (Novagen), from plasmids pDZ19 and pDZ12, respectively. TRA-1(Z-5) and TRA-1(Z-2) contain amino acids 6–382 and 6–279, respectively⁶. GLI, containing amino acids 210–1,107 (ref. 4), was purified using the PinPoint Xa protein purification system (Promega). Protein concentrations were determined by Bradford analysis.

Oligonucleotide sequences. *EBG-9*: TGGACGATTAGATATGAGATGATAA GAAATTAATATGAGTAGATATGAGTAGATAAGAAATTAATAATGAAATG GAAATGTGCGCCCTATAGTGAGTCGTATTA. *EBG-11*: TGGACGATTATGA AATGGAAATGTACAAATAATAGAAACGAAATAGTAAGAAATGAAATTT TGGAAACCAAAATCTCGCCCTATAGTGAGTCGTATTA. *SPS-1*: TGAAAAA GGAAACAGACATTTTCGCCCTATAGTGAGTCGTATTA. *SPS-2*: TTTATTAA CAAGAAAAACAAATTCGTAAAGATTACAAGTTCGCCCTATAGTTGAGTC GTATTA. *SPS-3*: ATGAAATCGAAATGGTACAAATAATAGAAACGAAATG AGTAAGAAATGAAATTTTGGAAACCAAAATCTCGCCCTATAGTTGAGTC GTATTA. *SPS-4*: TTTATTAAACAAAGAAAAACAAATTCGTAAAGATTACAA GTTGAGGTCGAGTGGACGATTTTCGCCCTATAGTGAGTCGTATTA. *SPS-7*: CGAAATGAGTAAGAAATGAAATTTTCGCCCTATAGTTGAGTCGTATTA. *SPS-9*: CGAAATGAGTAAGAAATTTTGGAAACCAAAATTCGTAA A. GGAAACAGACATTTTCGCCCTATAGTTGAGTCGTATTA. *SPS-11*: GTAGATATGAGTAGATAAGAAATTAATAATGAAATGGAAATGTACAAA TAATAGAAATTCGCCCTATAGTTGAGTCGTATTA.

Received 29 January; accepted 23 April 1999.

- Weed, M., Mundlos, S. & Olsen, B. R. The role of sonic hedgehog in vertebrate development. *Matrix Biol.* **16**, 53–58 (1997).
- Hodgkin, J. Sex determination in the nematode *C. elegans*: analysis of *tra-3* suppressors and characterization of *fem* genes. *Genetics* **114**, 15–52 (1986).
- Hodgkin, J. A genetic analysis of the sex-determining gene, *tra-1*, in the nematode *Caenorhabditis elegans*. *Genes Dev.* **1**, 731–745 (1987).
- Kinzel, K. W., Ruppert, J. M., Bigner, S. H. & Vogelstein, B. The GLI gene is a member of the Kruppel family of zinc-finger proteins. *Nature* **332**, 371–374 (1988).
- Alexandre, C., Jacinto, A. & Ingham, P. W. Transcriptional activation of hedgehog target genes in *Drosophila* is mediated directly by the cubitus interruptus protein, a member of the GLI family of zinc finger DNA-binding proteins. *Genes Dev.* **10**, 2003–2013 (1996).
- Zarkower, D. & Hodgkin, J. Molecular analysis of the *C. elegans* sex-determining gene *tra-1*: a gene encoding two zinc finger proteins. *Cell* **70**, 237–249 (1992).
- Kuwabara, P. E., Okkema, P. G. & Kimble, J. *tra-2* encodes a membrane protein and may mediate cell communication in the *Caenorhabditis elegans* sex determination pathway. *Mol. Biol. Cell* **3**, 461–473 (1992).
- Hodgkin, J. A. & Brenner, S. Mutations causing transformation of sexual phenotype in the nematode *Caenorhabditis elegans*. *Genetics* **86**, 275–287 (1977).
- Okkema, P. G. & Kimble, J. Molecular analysis of *tra-2*, a sex determining gene in *C. elegans*. *EMBO J.* **10**, 171–176 (1991).
- Spieth, J., Brooke, G., Kuersten, S., Lea, K. & Blumenthal, T. Operons in *C. elegans*: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. *Cell* **73**, 521–532 (1993).
- Goodwin, E. B., Okkema, P. G., Evans, T. C. & Kimble, J. Translational regulation of *tra-2* by its 3' untranslated region controls sexual identity in *C. elegans*. *Cell* **75**, 329–339 (1993).
- Zarkower, D. & Hodgkin, J. Zinc fingers in sex determination: only one of the two *C. elegans* Tra-1 proteins binds DNA in vitro. *Nucleic Acids Res.* **21**, 3691–3698 (1993).
- Stutz, F. & Rosbash, M. Nuclear RNA export. *Genes Dev.* **12**, 3303–3319 (1998).
- Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
- Seydoux, G. & Fire, A. Soma–germline asymmetry in the distributions of embryonic RNAs in *Caenorhabditis elegans*. *Development* **120**, 2823–2834 (1994).
- Kuwabara, P. E. & Kimble, J. Molecular genetics of sex determination in *C. elegans*. *Trends Genet.* **8**, 164–168 (1992).
- Kuwabara, P. E., Okkema, P. G. & Kimble, J. Germ-line regulation of the *Caenorhabditis elegans* sex-determining gene *tra-2*. *Dev. Biol.* **204**, 251–262 (1998).
- Jan, E., Yoon, J. W., Walterhouse, D., Iannaccone, P. & Goodwin, E. B. Conservation of the *C. elegans* *tra-2* 3'UTR translational control. *EMBO J.* **16**, 6301–6313 (1997).
- Goodwin, E. B., Hofstra, K., Hurney, C. A., Mango, S. & Kimble, J. A genetic pathway for regulation of *tra-2* translation. *Development* **124**, 749–758 (1997).
- Mello, C. & Fire, A. DNA transformation. *Methods Cell Biol.* **48**, 451–482 (1995).
- Fire, A. Histochemical techniques for locating *Escherichia coli* beta-galactosidase activity in transgenic organisms. *Genet. Anal. Tech. Appl.* **9**, 151–158 (1992).
- Hodgkin, J. Molecular cloning and duplication of the nematode sex-determining gene *tra-1*. *Genetics* **133**, 543–560 (1993).

Acknowledgements. We thank D. Zarkower for plasmids; R. Goldman, R. Holmgren and R. Lamb for use of their confocal microscopes; J. Austin, A. Bejsovec, E. Haag, and P. Okkema for critical reading of the manuscript; J. W. Yoon for GLI protein; and S. Adam, J. Kramer, S. Huang, and members of the Goodwin laboratory for discussions. We acknowledge the *Caenorhabditis elegans* Genetics Center for worm strains. This work was supported by the NIH, and by the Chicago Just Cause Organization to E.B.G. and the Chicago Baseball Cancer Charities grant to L.E.G.

Correspondence and requests for materials should be addressed to E.B.G. (e-mail: eb778@nwu.edu).

The tyrosine kinase c-Abl regulates p73 in apoptotic response to cisplatin-induced DNA damage

JianGen Gong^{*,†}, Antonio Costanzo^{†,‡}, Hong-Qiong Yang^{*}, Gerry Melino[§], William G. Kaelin Jr^{||}, Massimo Levrero[‡] & Jean Y. J. Wang^{*}

^{*} Department of Biology and the Cancer Center, University of California, San Diego, La Jolla, California 92093-0322, USA

[‡] Laboratory of Gene Expression, Fondazione A Cesalpino, University of Rome "La Sapienza", Viale del Policlinico 155, 00161 Rome, Italy

[§] Biochemistry Laboratory, Istituto Dermopatico dell'Immacolata-IRCCS, Department of Experimental Medicine, University of Rome "Tor Vergata" 00133, Rome, Italy

^{||} Howard Hughes Medical Institute, and Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA

[†] These authors contributed equally to this work

Cancer chemotherapeutic agents such as cisplatin exert their cytotoxic effect by inducing DNA damage and activating programmed cell death (apoptosis). The tumour-suppressor protein p53 is an important activator of apoptosis. Although p53-deficient cancer cells are less responsive to chemotherapy, their resistance is not complete, which suggests that other apoptotic pathways may exist. A p53-related gene, *p73*, which encodes several proteins as a result of alternative splicing^{1,2}, can also induce apoptosis³. Here we show that the amount of p73 protein in the cell is increased by cisplatin. This induction of p73 is not seen in cells unable to carry out mismatch repair and in which the nuclear enzyme c-Abl tyrosine kinase is not activated by cisplatin. The half-life of p73 is prolonged by cisplatin and by co-expression with c-Abl tyrosine kinase; the apoptosis-inducing function of p73 is also enhanced by the c-Abl kinase. Mouse embryo fibroblasts deficient in mismatch repair or in c-Abl do not upregulate p73 and are more resistant to killing by cisplatin. Our results indicate that c-Abl and p73 are components of a mismatch-repair-dependent apoptosis pathway which contributes to cisplatin-induced cytotoxicity.

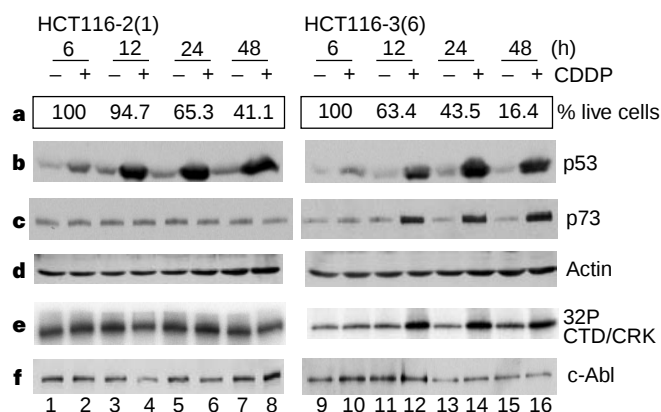


Figure 1 Induction of p73 by cisplatin (CDDP) correlated with activation of c-Abl tyrosine kinase and increased cell death. HCT116-1 cells (lanes 1–8) are defective in mismatch repair, but HCT116-3(6) cells (lanes 9–16) are proficient. **a**, Time course of cell death in 25 μ M cisplatin (see Methods); **b**, anti-p53 immunoblot; **c**, anti-p73 immunoblot; **d**, anti-actin immunoblot; **e**, phosphorylation (³²P-labelling) of CTD/CRK by c-Abl tyrosine kinase; **f**, anti-Abl immunoblot showing c-Abl protein in each kinase reaction. The c-Abl tyrosine kinase is active in both cell types, but it is only activated further by cisplatin in HCT116-3(6) cells¹¹.

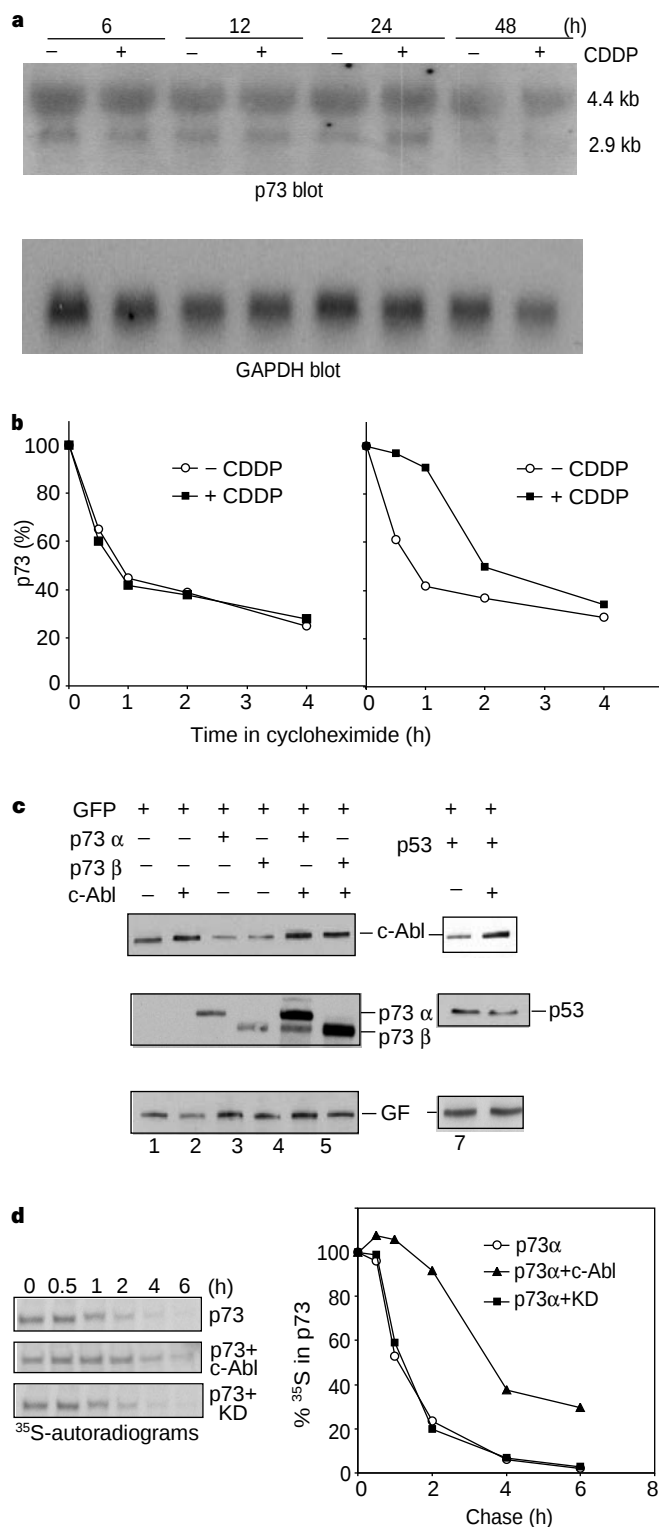


Figure 2 Stabilization of p73 protein by cisplatin and by c-Abl tyrosine kinase. **a**, Northern blots. The amount of p73 or GAPDH (glyceraldehyde-3-phosphate dehydrogenase, loading control) mRNA in 10 μ g total RNA was determined by hybridization with their cDNA probes. **b**, Cisplatin treatment prolonged the half-life of p73 protein in HCT116-3(6) (right) but not in HCT116-2(1) (left) cells (see Methods). **c**, Co-expression (see Methods) of c-Abl caused the accumulation of p73- α or - β but not of p53. The lower haemagglutinin (HA)-reactive band in lane 5 was a truncated form of p73- α ¹⁹. **d**, Co-expression of c-Abl, but not of a kinase-defective (KD) c-Abl mutant, increased the half-life of p73- α . Transfected Saos-2 cells were pulse-labelled and chased as described in Methods. p73 protein was immunoprecipitated with anti-HA antibodies and the amount of ³⁵S was quantified by Phosphorimaging.

To evaluate the role of p73 in DNA-damage-induced apoptosis, we chose cisplatin (*cis*-diam-minedichloroplatinum (II) as a damage inducer because it can kill most cell types. Cells defective in mismatch repair, such as those derived from the hereditary cancer syndrome human non-polypoid colon carcinoma, are most resistant to cisplatin⁴⁻⁷. We examined the regulation of p73 by cisplatin in a mismatch-repair-defective human cell line HCT116, which lacks the *MLH1* gene on chromosome 3. HCT116-3(6) cells have a functional *MLH1* gene as a result of transfer of chromosome 3 (ref. 8); the control cells, HCT116-2(1) cells, are also defective, but contain transferred chromosome 2 (ref. 8). Cisplatin kills HCT116-3(6) cells faster than HCT116-2(1) cells (Fig. 1a), consistent with previous reports⁸. This differential sensitivity to cisplatin could not be attributed to p53, which was induced in both cell types at a comparable rate and to a similar extent (Fig. 1b; see Fig. 1d for loading control). As treatment with cisplatin increased p73 protein in HCT116-3(6) cells (Fig. 1c, lanes 9–16) but not in HCT116-2(1) cells (Fig. 1c, lanes 1–8), the increased sensitivity of HCT116-3(6) cells to cisplatin could be correlated with the combined induction of p73 and p53. The induction of p73 by cisplatin was not due to an increase in the p73 messenger RNA (Fig. 2a); instead, the half-life of p73 was regulated by cisplatin. In untreated HCT116-2(1) and HCT116-3(6) cells, the half-life of p73 was ~45 min (Fig. 2b). After 24 hours of cisplatin treatment, the half-life of p73 increased to 2 hours in HCT116-3(6) cells, but remained unchanged in HCT116-2(1) cells (Fig. 2b).

Cisplatin can activate the nuclear c-Abl tyrosine kinase^{9,10}, but only in mismatch-repair-proficient cells¹¹ (compare lanes 1–8 to

lanes 9–16 of Fig. 1e, f). As the induction of p73 correlates with the activation of c-Abl tyrosine kinase (Fig. 1c, e, lanes 9–16), we tested the effect of c-Abl on the stability of p73. We co-expressed two isoforms of haemagglutinin-tagged human p73 protein (α and β)² with c-Abl in human osteosarcoma cells (Saos-2; Fig. 2c). This co-expression of c-Abl led to an increase in the amount of p73- α and p73- β (compare lanes 5, 6 to lanes 3, 4 in Fig. 2c). In contrast, co-expression with c-Abl did not increase the abundance of p53 (compare lanes 8 and 7 in Fig. 2c), or affect the levels of green fluorescent protein (GFP), which was used to normalize the transfection efficiency (Fig. 2c, lanes 1–8). To confirm that c-Abl can stabilize p73, transfected Saos-2 cells were pulse-labelled with ³⁵S-Met/Cys followed by a six-hour chase (Fig. 2d): the transiently expressed p73- α had a half-life of 30–40 min (Fig. 2d: p73 alone), which is comparable to that of the endogenous p73 in HCT116 cells (compare Fig. 2b, d). When p73- α was co-expressed with c-Abl, its half-life increased to 3.5 h (Fig. 2d: p73+c-Abl). The stabilization of p73 required the kinase activity of c-Abl, because co-expression with kinase-defective c-Abl (Fig. 2d: p73 + KD) did not affect the half-life of p73. The stabilization of p73 by the co-expression with c-Abl was also seen in several other human and mouse cell lines (for example, Fig. 3d). Together, the results shown in Figs 1 and 2 indicate that the activation of c-Abl tyrosine kinase contributes to the accumulation of p73 protein in cisplatin-treated cells.

Ectopic expression of a similar p73 protein induces apoptosis in p53-defective human Saos-2 cells³, as does ectopic expression of human p73- α or p73- β (A.C. and M.L., unpublished results). To

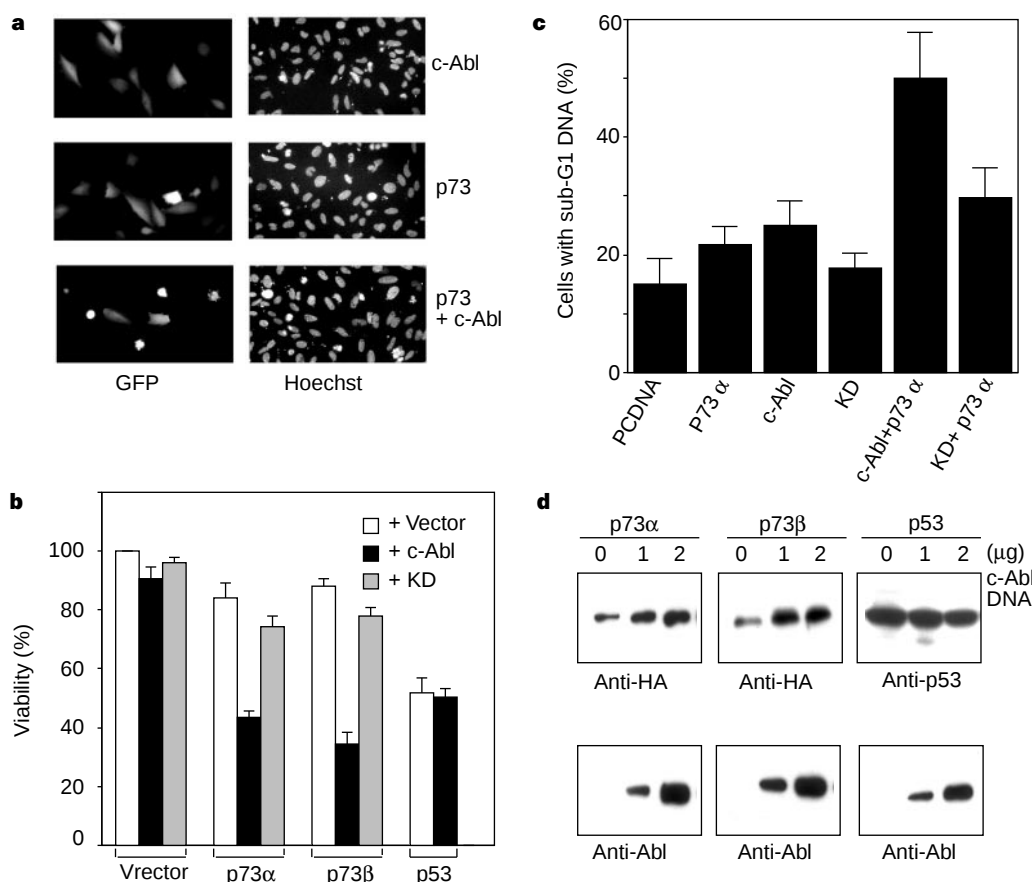


Figure 3 c-Abl enhances the apoptosis function of p73, but not p53. Abl-deficient 3T3 cells were co-transfected with the indicated expression plasmids plus the appropriate GFP expression plasmid, and dead GFP-positive cells counted (see Methods). **a**, Examples of GFP-positive cells and staining with Hoechst dye. **b**, Viable cells (green with diffuse Hoechst staining) were counted and normalized to the number in a vector-transfected sample. Values are means and standard deviations from four independent experiments, with at least 300 GFP-positive

cells counted per transfection. **c**, GFP-positive cells with sub-G1 DNA content were counted as described in Methods. Values are means and standard deviations from three independent experiments. Only the combined expression of p73 and c-Abl caused a significant increase in the sub-G1 DNA content over background (vector alone). **d**, Co-expression with c-Abl (anti-Abl blots) increased the amount of p73- α and p73- β (anti-HA blots) but not of p53 protein (anti-p53 blot) in Abl-deficient 3T3 cells.

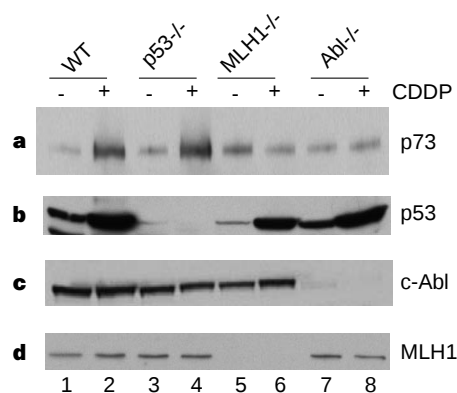


Figure 4 Differential induction of p53 and p73 by cisplatin in mouse embryo fibroblasts with defined mutations. Mouse embryo fibroblasts of the indicated genotypes were prepared as described in Methods. Total cellular proteins (300 μ g each) were extracted from control and cisplatin-treated cultures (25 μ M for 24 h) and the levels of p73, p53, c-Abl and MLH1 were determined by immunoblotting with specific antibodies. **a**, Anti-p73 immunoblot; **b**, anti-p53 immunoblot; **c**, anti-Abl immunoblot; **d**, anti-MLH1 immunoblot.

test the effect of c-Abl on the death-inducing function of p73, we transiently expressed human p73- α or - β in Abl-deficient mouse 3T3 cells (Fig. 3) and followed the fate of transfected cells by including an appropriate GFP-expression plasmid in the transfection (see Methods). Unlike Saos-2 cells, Abl-deficient 3T3 cells were resistant to ectopic expression of p73- α or - β (Fig. 3a–c: p73- α or p73- β plus vector). In contrast, Abl-deficient 3T3 cells were killed by expression of p53 (Fig. 3b, p53 plus vector). Expression of c-Abl alone did not affect cell viability either (Fig. 3a–c, c-Abl), indicating that the stabilization of endogenous p73 alone might not be sufficient to kill these 3T3 cells. However, co-expression of c-Abl and p73 caused cell death (Fig. 3a–c: p73- α or p73- β plus c-Abl), whereas the death inducing function of p53 was not affected by c-Abl (Fig. 3b: p53 plus c-Abl). A kinase-defective c-Abl mutant was less effective at enhancing the death-inducing function of p73 under identical conditions (Fig. 3b, c: p73 plus KD). Co-expression with c-Abl also caused p73 to accumulate in these Abl-deficient 3T3 cells but did not affect p53 levels (Fig. 3d). Although c-Abl kinase stabilizes and activates p73, we did not detect any phosphotyrosine in p73 (results not shown), suggesting that other proteins could also be involved in c-Abl-mediated regulation of p73.

Our results indicate that cisplatin can induce two parallel death-response pathways, one dependent on p53 and the other on p73. To confirm this, we assayed for induction of p53 and p73 in mouse embryo fibroblasts (MEFs) derived from genetically defined mutants of p53 (ref. 12), c-Abl¹³ and MLH1 (ref. 14). The absence of p53, c-Abl or MLH1 in the respective mutant MEFs was verified by immunoblotting with specific antibodies against each protein (Fig. 4b–d). Exposure to cisplatin induced both p53 and p73 in wild-type MEFs (Fig. 4a, b, lanes 1, 2). p53 was induced in MLH1- and c-Abl-deficient MEFs (Fig. 4b, lanes 5–8), showing that mismatch repair and c-Abl tyrosine kinase are not essential for the regulation of p53. p73 was induced in p53-deficient cells (Fig. 4a, lanes 3, 4), consistent with the regulation of p73 being p53-independent. Although p73 was expressed in MLH1- and c-Abl-deficient MEFs, treatment with cisplatin did not increase p73 in the absence of MLH1 or c-Abl function (Fig. 4a, lanes 5–8), indicating that MLH1 and c-Abl are required for induction of p73 by cisplatin.

We assessed the role of p73 in cell death by comparing the sensitivity of wild-type and mutant MEFs to killing by cisplatin (Fig. 5a, b). MEFs lacking p53 were less sensitive to cisplatin in that they died more slowly (Fig. 5a, p53^{-/-}) and survived better than wild-type MEFs in lower concentrations of cisplatin (Fig. 5b, p53^{-/-}). The killing of p53-deficient MEFs by cisplatin could be due to induction of p73 in these cells (Fig. 4a). c-Abl- or MLH1-

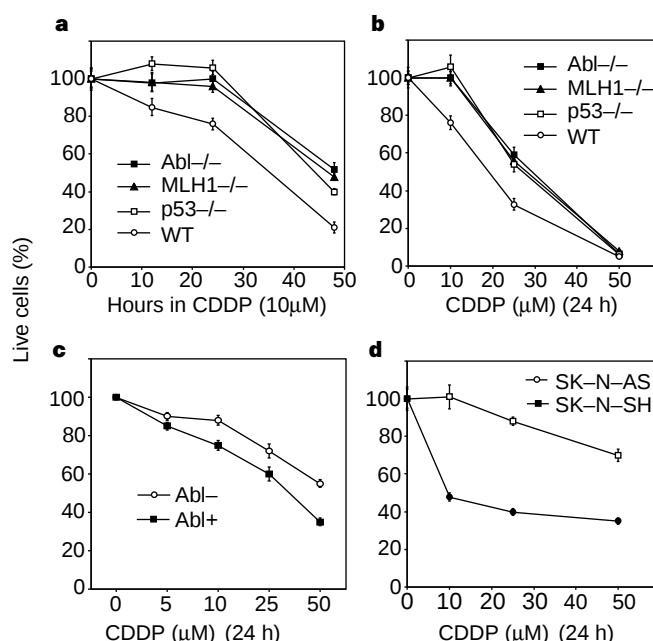


Figure 5 Defective cellular MLH1, c-Abl or p73 reduces the sensitivity to cisplatin. **a**, Kinetics of killing the indicated MEFs by 10 μ M cisplatin. **b**, Dose response of the indicated MEFs to cisplatin 24 h after drug addition. **c**, Re-introduction of c-Abl-sensitized Abl-deficient cells to attack by cisplatin. **d**, Dose response of two neuroblastoma cell lines to cisplatin 24 h after drug addition. Both cell lines expressed wild-type p53, SK-N-SH cells expressed p73 (ref. 1) but SK-N-AS cells did not; p73-negative cells were less sensitive to killing by cisplatin.

deficient MEFs also died more slowly (Fig. 5a) and survived lower concentrations of cisplatin better than their wild-type counterparts (Fig. 5b). The reduced sensitivity of c-Abl- and MLH1-deficient MEFs may be related to the lack of p73 accumulation (Fig. 4a). We also compared the sensitivity to cisplatin of Abl-deficient 3T3 cells before and after re-expression of c-Abl and found that reintroducing c-Abl increased the sensitivity to cisplatin (Fig. 5c). We also used two human neuroblastoma cell lines, SK-N-AS and SK-N-SH, to compare sensitivity to cisplatin. These cell lines, which are independently derived and not related to each other, were selected as they both express wild-type p53, but only SK-N-SH cells express p73 (ref. 1); both are sensitive to killing after transfection with human p73 (G.M., unpublished results), indicating that the apoptotic pathway downstream of p73 is intact. Consistent with our model, the p73-negative SK-N-AS cells are more resistant to cisplatin than the p73-positive SK-N-SH cells (Fig. 5d). These results together show that p73 induction as well as p53 accumulation contribute to cisplatin-induced cytotoxicity.

Our results show that p73 and p53 are regulated by distinct mechanisms: p53 is induced by a variety of DNA-damaging agents^{9,15}, whereas p73 is induced by cisplatin (Figs 1, 4) but not by ultraviolet radiation¹. c-Abl tyrosine kinase is also activated by cisplatin but not by ultraviolet light⁹, consistent with a link between c-Abl and p73. Activation of c-Abl and p73 by cisplatin requires a functional MLH1 gene, but induction of p53 does not, so defective c-Abl activation and p73 induction may partly account for the resistance of MLH1-deficient cells to cisplatin. In MSH2-deficient mice, which are also defective in mismatch repair, the intestinal crypt cells are resistant to cisplatin⁷; analysis of crypt cells in MSH2- and p53-mutant mice has revealed a p53-independent apoptosis response that is dependent on MSH2 (ref. 7). In testicular teratocarcinoma cells, which are sensitive to cisplatin, there is a p53-independent death response to cisplatin¹⁶. The induction of p73 by cisplatin described here might explain this p53-independent cell death, as activation of p53 alone is evidently not sufficient to

account for all the cytotoxic effects of cisplatin. Modulators of this p53-independent cell-death pathway, which involves mismatch repair, c-Abl and p73, may be able to increase the response of cancer cells to cisplatin. □

Methods

Cell culture and transfection. Human colon-carcinoma lines HCT116-2(1) and HCT116-3(6) (ref. 8), osteosarcoma line Saos-2, Abl-deficient 3T3 cells, or primary MEFs were cultured in standard medium plus fetal bovine serum. Retroviral-mediated gene transfer was used to reconstitute c-Abl expression in Abl-deficient 3T3 cells¹⁷ and polyclonal populations of reconstituted cells were used. *Abl*^{-/-}, *p53*^{-/-} and *MLH1*^{-/-} MEFs were prepared by mating the appropriate heterozygous mutant mice¹²⁻¹⁴. MEFs were prepared between day 11–13 of pregnancy and genotyped by PCR of embryo DNA¹²⁻¹⁴. Transient transfections were performed either with Superfect reagent (Gibco-BRL) or by using calcium phosphate precipitation¹⁸.

Immunoblotting. Endogenous p73 protein was detected by immunoblotting with a monoclonal anti-p73 antibody (ER15), as described¹⁹. The following antibodies were used: anti-HA (12CA5; Boehringer Mannheim), anti-Abl (8E9; Pharmingen), anti-p53 PAb421 (Oncogene Science), anti-p53 C-19 (Santa Cruz Biotech), anti-MLH1 (Ab-2; Oncogene Science), anti-GFP (Clontech), anti-actin (Oncogene Science). To prepare total proteins, cells were extracted with lysis buffer (50 mM Tris, pH 8, 120 mM NaCl and 0.5% NP-40) and the insoluble pellet discarded after centrifugation. The protein concentration in the soluble fractions was determined by using a BioRad dye-binding assay.

c-Abl tyrosine kinase assay. c-Abl tyrosine kinase activity was measured in an immune-complex kinase assay as described²⁰. A fusion protein (CTD–CRK), containing a modified C-terminal-repeat domain of RNA polymerase II (CTD) and the Abl phosphorylation site in the adapter protein CRK (J.G. and J.Y.J.W., unpublished), was used as a specific substrate to assay c-Abl kinase activity.

Measurement of p73 half-life. Decay of p73 protein in cycloheximide (Fig. 2b): cycloheximide (20 µg ml⁻¹) was added to cells 24 h after cisplatin (25 µM) addition and the cellular levels of p73 or actin at different time points were determined by immunoblotting; the amount of p73 relative to actin was normalized to that at the zero-time point. ³⁵S-pulse chase (Fig. 2d): Saos-2 cells were transfected with 3 µg each of pCDNA-HAP73-α with or without 3 µg pMSCV-c-Abl, or with a kinase-defective c-Abl (KD) containing a point mutation in the lysine residue critical for ATP binding. At 24 h post-transfection, cells were labelled with 250 µCi ml⁻¹ of Translabel (containing ³⁵S-Met and ³⁵S-Cys; Amersham) for 60 min. Unlabelled methionine and cysteine (1 mg ml⁻¹) were added and the cells collected at the indicated times; the ³⁵S-labelled p73-α in the anti-HA immunoprecipitate from each time point was quantified by Phosphorimager and normalized to that of the zero time point.

Cell-death assays. The killing of HCT116 cells by cisplatin (25 µM) was assayed by counting live cells, based on their trypan-blue exclusion. The percentage of live cells was determined by comparing live cells in a cisplatin-treated culture with those in an untreated culture collected at the same time (Fig. 1a). The killing of transfected Abl-deficient 3T3 cells (Fig. 3) was determined by chromatin condensation (Fig. 3a, b) or sub-G1 DNA content (Fig. 3c). In the chromatin-condensation assay, Abl-deficient 3T3 were co-transfected with 0.5 µg pEGFP (Clontech) and 3 µg each of pCDNA3-HA (vector), pCDNA-HAP73-α, pCDNA-HAP73-β or pCDNA3-p53 plasmid, either in the presence or absence of 3 µg MSCV-c-Abl or pMSCV-c-Abl-KD expression vector. At 96 h after transfection, cells were stained with Hoechst 33278 dye. The percentage of viable cells (diffuse Hoechst staining of green cells) after each transfection was determined and normalized to that of the pCDNA3-HA vector-transfected culture. In a fluorescence-activated cell-sorting (FACS) assay, c-Abl-deficient 3T3 cells were co-transfected with 4 µg of the indicated expression plasmids and 1.5 µg of a plasmid encoding a membrane-localized GFP²¹. At 48 h post-transfection, cells were permeabilized in PBS plus 0.1% Triton X-100 and stained with propidium iodide²². The DNA content of GFP-positive cells was determined with a FACScan and Cellfit program. The sensitivity of c-Abl-deficient and c-Abl-reconstituted 3T3 cells to cisplatin (Fig. 5c) was determined by Hoechst staining. The sensitivity of MEFs to cisplatin (Fig. 5a, b) was determined by counting the total number of cells using crystal violet stain. MEFs were seeded at 4 × 10⁴ per well into 24-well plates and treated with cisplatin one day later. At the indicated time after cisplatin

addition, cells were fixed and stained with 0.1% crystal violet. After extensive washing with PBS, the stain was extracted with 10% acetic acid and the dye concentration was determined from its absorbance at 600 nm. The optical density units of cisplatin-treated wells (in triplicate) at each time point were presented as a percentage of the units recovered from untreated wells (also in triplicate) collected at the same time point, with standard deviations (Fig. 5a, b).

Received 21 April; accepted 6 May 1999.

- Kaghad, M. *et al.* Monoallelically expressed gene related to p53 at 1-36, a region frequently deleted in neuroblastoma and other human cancers. *Cell* **90**, 809–819 (1997).
- De Laurenzi, V. *et al.* Two new p73 splice variants, gamma and delta, with different transcriptional activity. *J. Exp. Med.* **188**, 1–6 (1998).
- Jost, C., Marin, M. & Kaelin, W. p73 is a human p53-related protein that can induce apoptosis. *Nature* **389**, 191–194 (1997).
- Drummond, J., Anthony, A., Brown, R. & Modrich, P. Cisplatin and Adriamycin resistance are associated with MutL-alpha and mismatch repair deficiency in an ovarian cancer cell line. *J. Biol. Chem.* **271**, 19645–19648 (1996).
- Fink, D. *et al.* *In vitro* and *in vivo* resistance to cisplatin in cells that have lost DNA mismatch repair. *Cancer Res.* **57**, 1841–1845 (1995).
- Buermeier, A. B., Patten, C. W., Baker, S. M. & Liskay, R. M. The human MLH1 cDNA complements DNA mismatch repair defects in Mlh1-deficient mouse embryonic fibroblasts. *Cancer Res.* **59**, 538–541 (1999).
- Toft, N. J. *et al.* Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. *Proc. Natl Acad. Sci. USA* **96**, 3911–3915 (1999).
- Boyer, J. *et al.* Microsatellite instability, mismatch repair-deficiency, and genetic defects in human cancer cell lines. *Cancer Res.* **55**, 6063–6070 (1995).
- Liu, Z. *et al.* Three distinct signalling responses by murine fibroblasts to genotoxic stress. *Nature* **384**, 273–276 (1996).
- Kharbanda, S. *et al.* Activation of the c-Abl tyrosine kinase in the stress response to DNA-damaging agents. *Nature* **376**, 785–788 (1995).
- Nehme, A. *et al.* Differential Induction of c-Jun NH2-terminal kinase and c-Abl kinase in DNA mismatch repair-proficient and deficient cells exposed to cisplatin. *Cancer Res.* **57**, 3253–3257 (1997).
- Lowe, S. W., Schmitt, E. M., Smith, S. W., Osborne, B. A. & Jacks, T. p53 is required for radiation-induced apoptosis in mouse thymocytes. *Nature* **362**, 847–849 (1993).
- Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. Neonatal lethality and lymphopenia in mice with a homozygous disruption of the *c-abl* proto-oncogene. *Cell* **65**, 1153–1163 (1991).
- Edelmann, W. *et al.* Meiotic pachytene arrest in MLH1-deficient mice. *Cell* **85**, 1125–1134 (1996).
- Nelson, W. G. & Kastan, M. B. DNA strand breaks: the DNA template alterations that trigger p53-dependent DNA damage response pathways. *Mol. Cell. Biol.* **14**, 1815–1823 (1994).
- Zamble, D., Jacks, T. & Lippard, S. p53-dependent and -independent responses to cisplatin in mouse testicular teratocarcinoma cells. *Proc. Natl Acad. Sci. USA* **95**, 6163–6168 (1998).
- Pear, W. S., Nolan, G. P., Scott, M. L. & Baltimore, D. Production of higher-titer helper-free retrovirus by transient transfection. *Proc. Natl Acad. Sci. USA* **90**, 8392–8396 (1993).
- Chen, C. & Okayama, H. High efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745–2752 (1987).
- Marin, M. *et al.* Viral oncoproteins discriminate between p53 and the p53 homolog p73. *Mol. Cell. Biol.* **18**, 6316–6324 (1998).
- Baskaran, R. *et al.* Ataxia telangiectasia mutant protein activates c-Abl tyrosine kinase in response to ionizing radiation. *Nature* **387**, 516–519 (1997).
- Kaleja, R., Shen, T. & Beavis, A. Use of a membrane-localized green fluorescent protein allows simultaneous identification of transfected cells and cell cycle analysis by flow cytometry. *Cytometry* **29**, 286–291 (1997).
- Nicoletti, L., Migliorati, G., Pagliacci, M. C., Grignani, F. & Riccardi, C. A rapid and simple method for measuring thymocyte apoptosis by propidium iodide staining and flow cytometry. *J. Immunol. Meth.* **139**, 271–279 (1991).

Acknowledgements. We thank I. Hunton for the preparation of p53- and c-Abl-deficient MEFs, R. Kucherlapati and R. Kolodner for the MLH1-deficient MEFs, and R. Kolodner for critically reading the manuscript. This work was supported by grants from AIRC and CNR-PF Biotechnology (to M.L.) and from the NIH (to J.Y.J.W.).

Correspondence and requests for materials should be addressed to J.Y.J.W. (e-mail: jywang@ucsd.edu).

Interaction of c-Abl and p73α and their collaboration to induce apoptosis

Reuven Agami*, Giovanni Blandino†, Moshe Oren† & Yosef Shaul*

Departments of *Molecular Genetics and †Molecular Cell Biology, The Weizmann Institute of Science, Rehovot 76100, Israel

c-Abl, a non-receptor tyrosine kinase, is activated by agents that damage DNA. This activation results in either arrest of the cell cycle in phase G1 or apoptotic cell death, both of which are dependent on the kinase activity of c-Abl¹. p73, a member of the p53 family of tumour-suppressor proteins^{2,3}, can also induce apoptosis³. Here we show that the apoptotic activity of p73α requires the presence of functional, kinase-competent c-Abl. Furthermore, p73 and c-Abl can associate with each other, and

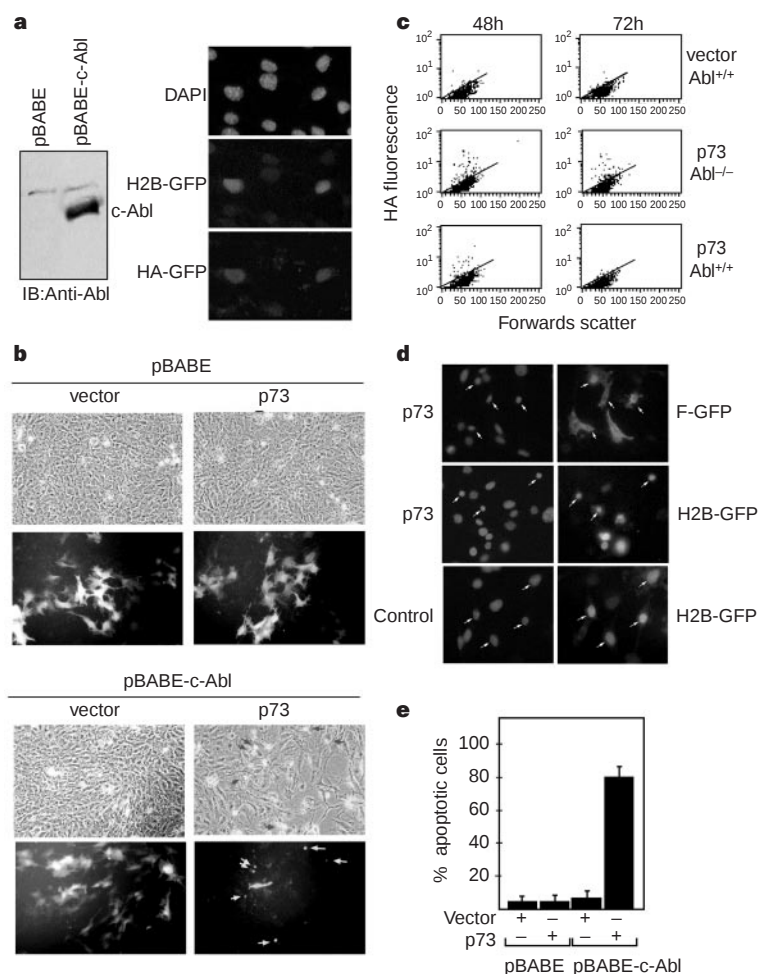


Figure 1 p73 induces apoptosis in a c-Abl-dependent manner. **a**, Abl^{-/-} fibroblasts, infected with either a pBabe-Puro retrovirus or a pBabe-Puro-c-Abl retrovirus and selected with puromycin. Cell extracts were separated on 7.5% SDS-PAGE and immunoblotted with the 8E9 anti-Abl antibody. Also shown are pBabe cells co-transfected with H2B-GFP (GFP fused to histone 2B) and HA-p73. Cells were stained with DAPI to visualize nuclei and with anti-HA antibody to visualize p73. The transfected positive cells also express p73. **b**, Cells were transfected with 3 µg of either a control vector of HA-p73 expression plasmid with 0.5 µg pEGFP DNA (a commercial expression plasmid encoding GFP) using lipofectAMINE (Gibco-BRL). Cells were photographed 96 h post-transfection under UV (bottom) or phase microscope (top). Cells with apoptotic morphology are indicated by arrowheads. **c**, Kinetics of disappearance of HA-p73α-positive cells. Cultures of pBabe-Puro-c-Abl cells (top and bottom) or pBabe-Puro cells

(middle) were transfected transiently with either empty vector (top) or HA-p73α expression plasmid (middle and bottom). At the indicated times after transfection, cells were fixed, stained with anti-HA antibody and analysed by FACS. The diagonal line represents an arbitrary division between HA-p73-positive cells (above line) and the rest of the cells in the culture. There is some nonspecific background staining (upper panels). **d**, pBabe-Puro-c-Abl cells growing on coverslips were transfected with either F-GFP (membrane-targeted GFP) or H2B-GFP (nucleus-associated GFP) with either p73 expression plasmid or the empty vector (control). Left, DAPI-stained fields; right, GFP fluorescence of the same fields. Note the shrunken condensed nuclei, indicating apoptotic morphology, in p73-transfected panels. **e**, Average percentage of apoptotic cells from five randomly chosen fields, with s.d.

this binding is mediated by a PxxP motif in p73 and the SH3 domain of c-Abl. We find that p73 is a substrate of the c-Abl kinase and that the ability of c-Abl to phosphorylate p73 is markedly increased by γ-irradiation. Moreover, p73 is phosphorylated *in vivo* in response to ionizing radiation. These findings define a pro-apoptotic signalling pathway involving p73 and c-Abl.

We investigated the relationship between p73 and c-Abl in the context of apoptosis using c-Abl nullizygous (Abl^{-/-}) mouse fibroblasts, stably transfected with either a c-Abl expression plasmid or the corresponding empty vector. As expected, the correct c-Abl protein was expressed in the former but not in the latter transfectants (Fig. 1a, left). Each type of cell was transfected transiently with an expression plasmid encoding haemagglutinin (HA)-tagged simian p73. Successfully transfected cells, identified using a co-transfected green fluorescent protein (GFP) expression plasmid, produced readily detectable amounts of HA-p73 (Fig. 1a, right). We then inspected these GFP-positive cells under a fluorescence

microscope, without fixation, to assess the rate of apoptosis in the transfected cultures. Mouse fibroblasts usually appear flat and well attached, but they shrink and round up when undergoing apoptosis. When c-Abl^{-/-} cells carrying only the empty vector (Fig. 1b, pBabe) were transfected transiently with p73, most of the transfectants retained their normal appearance (Fig. 1b). Hence, p73 is an inefficient inducer of apoptosis in these cells. In contrast, over-expression of p73 in cells reconstituted with c-Abl (pBabe-c-Abl) resulted in marked apoptosis, reflected by extensive rounding and shrinkage of the GFP-positive cells (Fig. 1b, right). When the HA-p73-transfected cells were monitored by fluorescence-activated cell sorting (FACS), the number of HA-p73-positive cells dropped sharply with time only in the presence of c-Abl (Fig. 1c). This observation was further confirmed by the analysis of fixed cultures; many of the GFP-positive p73 transfectants had shrunken, condensed nuclei, typical of apoptotic cells (Fig. 1d: arrows, upper two panels). This was seen with two different types of co-transfected

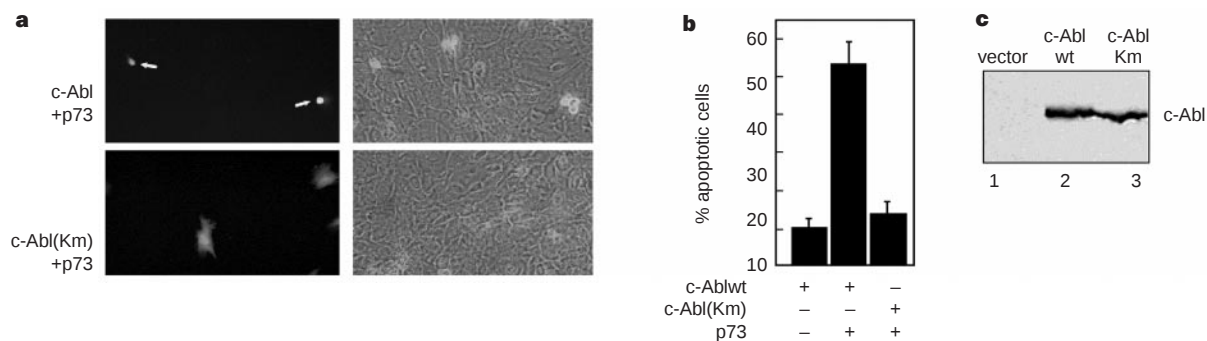


Figure 2 c-Abl and p73 collaborate to induce apoptosis in a kinase-dependent manner. **a**, Abl^{-/-} fibroblasts were transfected with 1.5 µg of each of the indicated expression plasmids, with 0.3 µg pEGFP. The c-Abl(Km) plasmid drives c-Abl synthesis with a point mutation in the kinase domain. This mutant is inactive in

kinase reactions (data not shown). Cells were photographed 72 h after transfection (arrowheads: apoptotic morphology). **b**, Average percentage of apoptotic cells from five randomly chosen fields. **c**, Protein extracts from cells transfected with the indicated plasmids were immunoblotted with the 8E9 anti-Abl antibody.

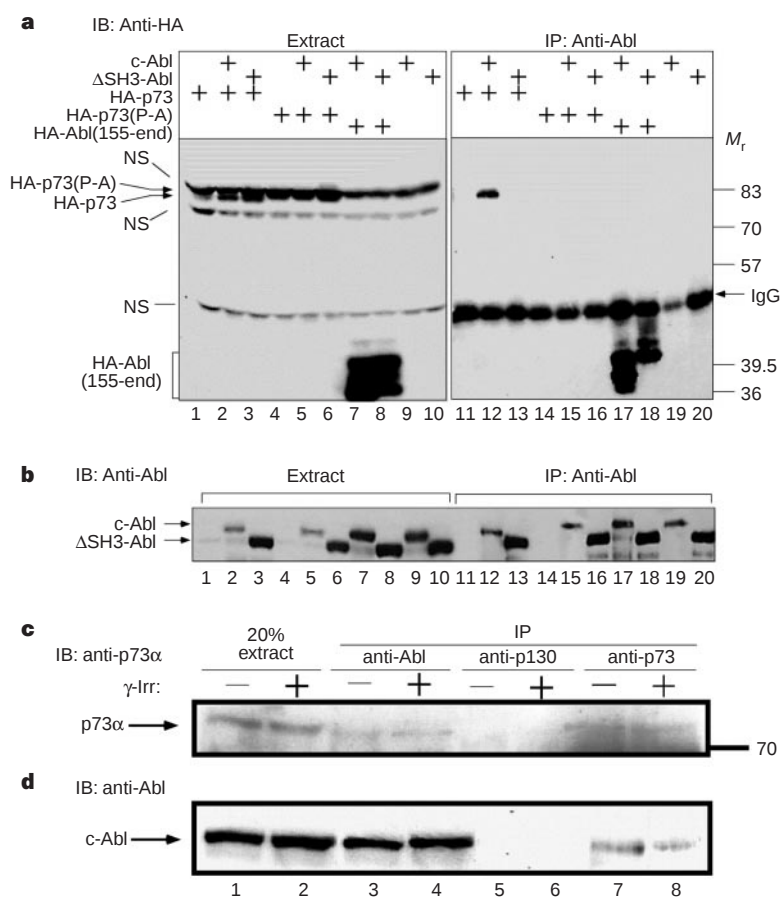


Figure 3 Interaction of p73 with c-Abl is mediated by the SH3 domain of c-Abl and a PxxP motif in p73. **a**, Cellular extracts were prepared from 293T cells transfected with 10 µg of the indicated expression plasmids. 90% of the extracts was immunoprecipitated with the K-12 anti-Abl antibody. The beads were extensively washed and bound proteins were eluted. 10% of the extracts (extract) and 50% of the eluted fractions (IP) were separated by 7.5% SDS-PAGE and immunoblotted with anti-HA antibody. The exposure times for the extract panel and the IP panel were 20 s and 5 min, respectively. The positions of HA-p73, its point mutant HA-p73(P-A), the c-Abl binding protein HA-Abl1(155-end) and a non-specific (NS)

background band are indicated. The P-A mutant migrates more slowly than the wild type. **b**, The same blot, re-probed with the 8E9 anti-Abl antibody. The positions of intact c-Abl and its mutant lacking the SH3 domain (ΔSH3-Abl) are shown. **c**, MCF-7 cells were either exposed to 20 Gy of γ-irradiation (γ-Irr) or left untreated. Two hours later, cells were lysed and cleared lysates were immunoprecipitated (IP) with the indicated antibodies. The eluted proteins and 20% of the total extracts were separated by 7.5% SDS-PAGE and immunoblotted with anti-p73α antibody. **d**, The blot in **c**, re-probed with the 8E9 anti-Abl antibody.

GFP (Fig. 1d, F-GFP and H2B-GFP). In contrast, nuclear condensation was not seen in cells transfected transiently with GFP alone (Fig. 1d, control). The data are plotted in Fig. 1e.

Our results indicate that p73 and c-Abl collaborate to induce apoptosis. This conclusion is further supported by experiments in which both p73 and c-Abl were co-transfected into c-Abl^{-/-} cells.

Extensive apoptosis was induced when p73 was introduced with wild-type c-Abl, but not with a kinase-defective c-Abl mutant (c-Abl(Km), Fig. 2a, b), even though the two types of c-Abl proteins were produced in comparable amounts in the transfected cultures (Fig. 2c). Thus, the kinase domain of c-Abl is required for the induction of apoptosis in collaboration with p73.

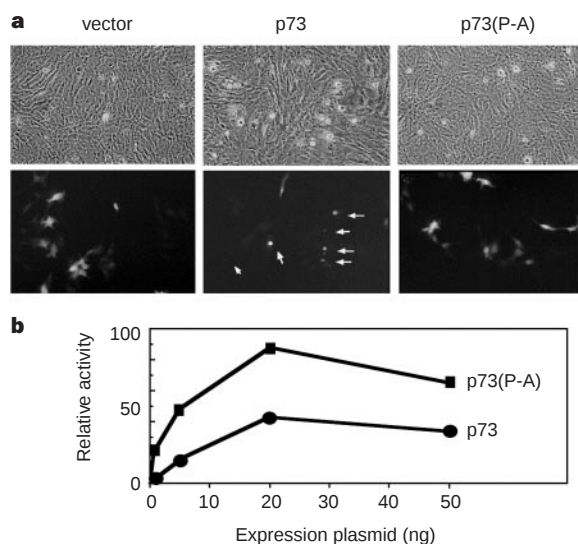


Figure 4 The PxxP motif of p73 is necessary for apoptosis but not for transcriptional activation. **a**, The c-Abl-reconstituted cells described in Fig. 1 were transfected with 3 μ g of expression plasmid encoding either p73 or p73(P-A) with 0.5 μ g of pEGFP. Cells were photographed 96 h after transfection. **b**, Hep3B cells were co-transfected with two reporter plasmids: one containing firefly luciferase under the *mdm2* promoter and the other containing renilla luciferase under the SV40-promoter (SVRL, Promega), with the indicated amounts of either p73 or p73(P-A) expression plasmids. Cells were collected 36 h after transfection and analysed using the dual luciferase kit (Promega). Relative activity was calculated as $100 \times$ firefly-luciferase/renilla-luciferase. Similar results were obtained in three independent experiments.

We investigated whether p73 and c-Abl could associate with each other by using cells co-transfected transiently with expression plasmids for c-Abl and HA-p73. To facilitate the detection of protein-protein interactions, we used human 293 cells, which are more readily transfectable than mouse c-Abl^{-/-} fibroblasts. As shown in Fig. 3a, HA-p73 was brought down by a c-Abl-specific polyclonal antibody, but only when wild-type c-Abl was co-expressed in the same cells (lanes 11, 12, 19). The same was true for Abi1, a c-Abl binding protein⁴ serving as a positive control (lane 17). Hence, p73 and c-Abl can interact specifically in transfected cells.

Inspection of the p73 protein sequence revealed that the putative linker region, located between the DNA-binding domain and the predicted oligomerization domain, contains a consensus PxxP motif—a potential site of interaction with the SH3 domain of c-Abl^{5,6}. The existence of such an interaction was investigated using a c-Abl mutant carrying a deletion that removes the SH3 domain (c-Abl Δ SH3), and a p73 mutant carrying a substitution from proline to alanine at residue 338 (P338A). Despite the fact that both mutants were expressed in transfected cells at levels comparable to those of the respective wild-type proteins (Fig. 3a, b, extract), none produced stable p73–c-Abl complexes (Fig. 3a, lanes 13–16). In contrast, the Abi1 polypeptide was brought down with either wild-type or SH3-domain-deleted c-Abl protein (lanes 17, 18), confirming that this interaction is independent of the c-Abl SH3 domain^{4,7}. Hence, the specific interaction between c-Abl and p73 requires the SH3 domain of c-Abl and the PxxP motif of p73.

To study p73–c-Abl interactions in non-transfected cells, extracts of MCF7 cells were either γ -irradiated or left untreated, and analysed using co-immunoprecipitation. p73 α was brought down by antibodies against c-Abl (Fig. 3c, lanes 3, 4), but not by control antibodies against p130 (lanes 5, 6). Specific co-precipitation was also seen in the reciprocal experiment, in which c-Abl was brought down by antibodies against p73 (Fig. 3d, lanes 7, 8). This indicates that c-Abl and p73 α interact specifically under physiological conditions in non-overexpressing cells. Neither the total amount of p73

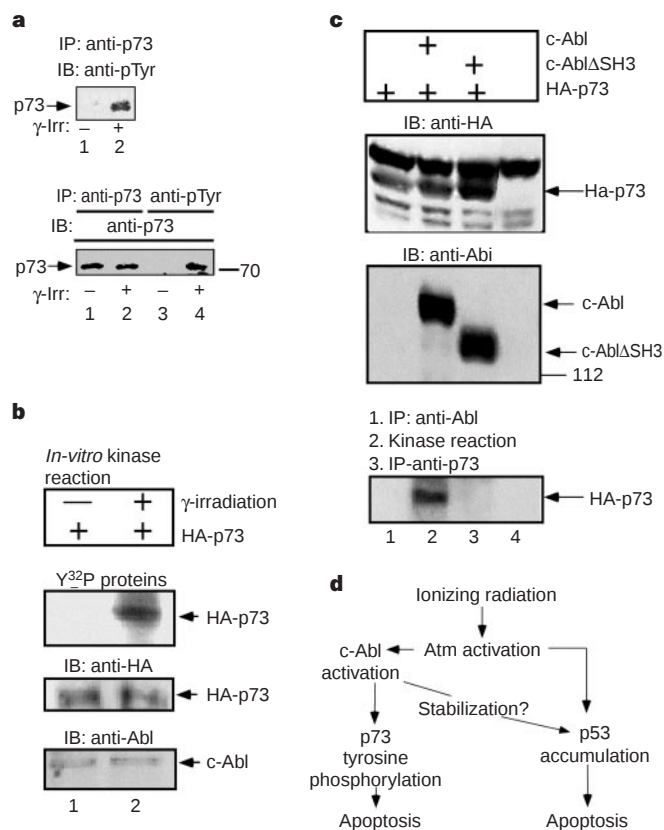


Figure 5 p73 is a substrate of the c-Abl kinase. **a**, MCF-7 cells were γ -irradiated and treated as in Fig. 3. Extracts were subjected to immunoprecipitation (IP) with anti-p73 α antibody and the precipitated proteins were analysed by western immunoblotting (IB) using anti-pTyr antibody. In a reciprocal experiment, the anti-p73 α antibody was used for IB (lower panel). **b**, MCF-7 nuclear extracts were prepared, either before or after γ -irradiation, and immunoprecipitated with the K-12 anti-Abl antibody. Immunoaffinity-purified HA-p73 protein, prepared from transiently transfected 293T cells, was added to the c-Abl-beads complex and incubated under kinase reaction conditions in the presence of [γ -³²P]ATP. The products were resolved by SDS-PAGE, blotted and treated with 1 M KOH for 2 h at 65 °C to detect solely phospho-tyrosines (upper panel). The quantity of c-Abl and HA-p73 in the reaction was monitored by the corresponding antibodies (IB, lower panels). **c**, 293T cells were transfected with the indicated expression plasmids and cellular protein extracts were prepared and analysed. Proper expression of the transfected plasmids was confirmed by immunoblotting with anti-HA and anti-Abl. Extracts were immunoprecipitated with the K-12 anti-Abl antibody and subjected to an *in vitro* kinase reaction for 1 h at 30 °C in the presence of [γ -³²P]ATP. The reaction was stopped and subjected to a second round of immunoprecipitation with the anti-HA antibody for 2 h at 4 °C. HA-p73 was eluted with 1 mg ml⁻¹ HA peptide. The eluates were resolved on 7.5% SDS-PAGE, blotted onto a nylon membrane and alkaline-treated as in **b**. **d**, Proposed signalling cascades triggered by ionizing radiation. Activation of the Atm kinase results in c-Abl kinase activation, leading to p73 tyrosine-phosphorylation and p53 accumulation. The accumulation of p53 is activated in part through direct phosphorylation of p53 by Atm^{12,22}. The activation of c-Abl may also contribute to p53 stabilization by protecting it against Mdm2-directed proteolysis²³.

nor its interaction with c-Abl were affected by γ -irradiation (γ -Irr, Fig. 3c, d).

To test whether the induction of apoptosis by c-Abl and p73 depends on their physical interaction, c-Abl-reconstituted fibroblasts (pBABE–c-Abl, Fig. 1) were transfected with either wild-type p73 or the P338A mutant of p73. In agreement with the data described earlier, wild-type p73 induced apoptosis in many of the transfected cells (Fig. 4a, p73). In contrast, the P338A mutant did not (Fig. 4a, p73(P-A)). This mutant could, however, support transcription of a luciferase reporter gene driven by the p53/p73-

responsive *mdm2* promoter. In fact, it was even more potent in this respect than intact wild-type p73 (Fig. 4b). In the case of the p53 protein, deletion of the PxxP motif also results in loss of apoptotic capacity^{8,9}. Importantly, although the deleted p53 can still activate many promoters, it fails to induce the promoter of the apoptosis-related *PIG3* gene¹⁰. Although the *PIG3* promoter does not appear to be a p73 target¹¹, p73(P-A) may therefore be defective for the activation of a limited subset of genes. Nevertheless, the overall transcriptional competence of p73(P-A) indicates that it retains many of its molecular interactions. Hence the defect in apoptosis is highly specific, implying a non-random relationship between the ability of p73 to bind c-Abl and its ability to induce apoptosis.

These findings raise the possibility that p73 is a substrate of the c-Abl kinase. The c-Abl kinase is activated by γ -irradiation¹, and endogenous p73 α becomes phosphorylated on tyrosine after irradiation (Fig. 5a). To determine whether c-Abl may be responsible for this phosphorylation, we immunoprecipitated endogenous c-Abl from MCF7 cells and measured its ability to phosphorylate purified HA-p73 *in vitro*. HA-p73 was tyrosine-phosphorylated only when the c-Abl was purified from irradiated cells (Fig. 5b). Moreover, overexpressed c-Abl from transfected 293T cells could efficiently phosphorylate the co-precipitated p73 α *in vitro* (Fig. 5c); as expected, this required the SH3 domain.

Collectively, our data support the existence of a signalling pathway that involves both p73 and c-Abl as mediators of apoptosis (Fig. 5d). Unlike p53, p73 protein levels do not increase following genotoxic stress². However, c-Abl does respond to at least certain conditions of genotoxic stress, including ionizing radiation. Under such conditions, the kinase activity of c-Abl is induced, presumably through the action of the stress-activated¹² ATM protein kinase^{13,14}. The facts that an intact kinase domain is essential for the cooperation of c-Abl with p73 in induction of apoptosis, and that p73 is a substrate of stress-induced c-Abl, indicate that p73 might be recruited into a stress-induced signalling complex through its association with c-Abl. □

Methods

Plasmid constructions and antibodies. To generate the pBABE-Puro-c-Abl plasmid, the retroviral vector pBABE-Puro was digested with *EcoRI* and *BamHI* and filled in by the Klenow enzyme. The c-abl fragment insert was excised from pSP6-c-Ablb (ref. 15) with *EcoRI* and *FspI*, and ligated to the pBABE-Puro vector after filling in. For transfection experiments, we constructed pSG5-c-Abl by cloning the *EcoRI*-*FspI* c-Ablb complementary DNA fragment into Bluescript KS (Stratagene) at the *EcoRI*-*SmaI* sites. The resulting plasmid was digested with *EcoRI* and *BamHI* and the c-Abl fragment was inserted into the pSG5 vector (Stratagene). We constructed the c-Abl-kinase-mutant expression plasmid pSG5-c-Abl(Km) using site-directed mutagenesis¹⁶ with the oligonucleotide 5'-GACGGTGGCCGTGCACACCTTGAAGGAG-3'. The c-Abl Δ SH3 expression vector was constructed by digesting pSG5-c-Abl with *HindIII* and *HincII* and re-ligating the deleted plasmid in the presence of a linker oligonucleotide.

The expression plasmid for wild-type simian p73 α (ref. 3) was a gift from W. J. Kaelin. The pSG5HA-p73 α PxxA mutant was prepared by PCR using the primers 5'-ATGAGCCACACAGGTGGGGAC-3', 5'-CCCCCTGCCGTGCGC-GCCCT-3' 5'-AGGCGCGCGACGGCAGGGGG-3' and 5'-GCTCTAGAT-CAGTGGATCTCGGCTCCGTG-3'. The resulting PCR product was digested with *XbaI* and *EcoRI* and inserted back into the HA-p73 cDNA to form HA-p73(P-A). We constructed the expression plasmid pSG5HA-Abl(155-end) by cloning the mouse cDNA coding for Abl(155-end) from total RNA of an Abl^{-/-} mouse cell line using RT-PCR (Promega) and the primers 5'-GCTCTAGAAGCGGAAGCCGAGAGAACAG-3' and 5'-CGGGATCCTAATCAGTATAGTGCATGA-3'. The resulting fragment was digested with *XbaI* and *BamHI*, and cloned into the pSG5-HA vector digested with *NheI* and *BamHI*. We used the pEGFP expression plasmid (Clontech) for expression of the GFP under the cytomegalovirus promoter/enhancer. A plasmid encoding a histone 2B/GFP fusion protein was a gift from T. Kanda¹⁷. The F-GFP plasmid has been described¹⁸. All constructions and mutants were verified by enzymatic digestion and sequencing.

The antibodies used for western-blot analysis are the monoclonals 12CA5 anti-HA, 8E9 anti-Abl (Pharmingen), and the polyclonal K-12 anti-Abl (sc-131), anti-phosphotyrosine PY-99 and anti-p73 α (Santa Cruz). The anti-HA monoclonal 11HA (Babco) was used for immunostaining of transfected cells¹⁹.

Cell transfection and retroviral infection. Cell transfection was done either by using the calcium phosphate co-precipitation method²⁰, or with lipofect-AMINE (Gibco-BRL) as recommended by the manufacturer. We generated high-titre retroviral supernatants by transfecting 293T cells with retroviral plasmids together with the Ψ -plasmid in the presence of 0.2 mM chloroquine, as described²¹. Cells were washed after 8 h and the medium was collected at intervals of 8–10 h for 38 h. Retroviral infection of the Abl^{-/-} cell line was done by incubation of a 50%-confluent 10-cm dish with 2 ml of retrovirus-containing medium for 4 h together with 10 μ g ml⁻¹ Polybrene. After a further 24 h, cells were selected with 2.5 μ g ml⁻¹ puromycin (Sigma).

Co-immunoprecipitation experiments. 293T cells were transfected with 20 μ g DNA per 10-cm dish, washed 12 h after transfection and collected after an additional 48 h. Cells were lysed in 300 μ l lysis buffer containing 10 mM β -glycerophosphate, 10 mM KCl, 1 mM EDTA, 1 mM EGTA, 2 mM MgCl₂, 0.5% N-P40, 1 M sucrose, 1 mM sodium vanadate, 1 mM DTT and a protease-inhibitor cocktail (Sigma), incubated for 10 min on ice and centrifuged at 13,000 r.p.m. for 10 min at 4 °C. Immunoprecipitation was carried out in 600 μ l lysis buffer containing 200 μ l cellular extract, 50 μ l of 50% slurry protein A-Sepharose beads (Pharmacia Tech) and 2 μ g of K-12 anti-Abl antibody for 2 h at 4 °C. The beads were then washed five times and the bound proteins were eluted by boiling in SDS-sample buffer and resolved by 7.5% SDS-PAGE.

Received 17 March; accepted 19 April 1999.

1. Kharbanda, S., Yuan, Z. M., Weichselbaum, R. & Kufe, D. Determination of cell fate by c-Abl activation in the response to DNA damage. *Oncogene* **17**, 3309–3318 (1998).
2. Kaghad, M. Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. *Cell* **90**, 809–819 (1997).
3. Jost, C. A., Marin, M. C. & Kaelin, W. J. p73 is a human p53-related protein that can induce apoptosis. *Nature* **389**, 191–194 (1997).
4. Shi, Y., Alin, K. & Goff, S. P. Abl-interactor-1, a novel SH3 protein binding to the carboxy-terminal portion of the Abl protein, suppresses v-abl transforming activity. *Genes Dev.* **9**, 2583–2597 (1995).
5. Rickles, R. J. et al. Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. *EMBO J.* **13**, 5598–5604 (1994).
6. Sparks, A. B., Quilliam, L. A., Thorn, J. M., Der, C. J. & Kay, B. K. Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries. *J. Biol. Chem.* **269**, 23853–23856 (1994).
7. Dai, Z. & Pendergast, A. M. Abi-2, a novel SH3-containing protein interacts with the c-Abl tyrosine kinase and modulates c-Abl transforming activity. *Genes Dev.* **9**, 2569–2582 (1995).
8. Walker, K. K. & Levine, A. J. Identification of a novel p53 functional domain that is necessary for efficient growth suppression. *Proc. Natl Acad. Sci. USA* **93**, 15335–15340 (1996).
9. Sakamuro, D., Sabbatini, P., White, E. & Pendergast, G. C. The polyproline region of p53 is required to activate apoptosis but not growth arrest. *Oncogene* **15**, 887–898 (1997).
10. Venot, C. et al. The requirement for the p53 proline-rich functional domain for mediation of apoptosis is correlated with specific PIG3 gene transactivation and with transcriptional repression. *EMBO J.* **17**, 4668–4679 (1998).
11. Zhu, J., Jiang, J., Zhou, W. & Chen, X. The potential tumor suppressor p73 differentially regulates cellular p53 target genes. *Cancer Res.* **58**, 5061–5065 (1998).
12. Banin, S. et al. Enhanced phosphorylation of p53 by ATM in response to DNA damage. *Science* **281**, 1674–1677 (1998).
13. Baskaran, R. et al. Ataxia telangiectasia mutant protein activates c-Abl tyrosine kinase in response to ionizing radiation. *Nature* **387**, 516–519 (1997).
14. Shafman, T. et al. Interaction between ATM protein and c-Abl in response to DNA damage. *Nature* **387**, 520–523 (1997).
15. Shivelman, E., Lifshitz, B., Gale, R. P., Roe, B. A. & Canaani, E. Alternative splicing of RNAs transcribed from the human abl gene and from the bcr-abl fused gene. *Cell* **47**, 277–284 (1986).
16. Kunkel, T. A., Roberts, J. D. & Zakour, R. A. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367–382 (1987).
17. Kanda, T., Sullivan, K. F. & Wahl, G. M. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr. Biol.* **8**, 377–385 (1998).
18. Jiang, W. & Hunter, T. Analysis of cell-cycle profiles in transfected cells using a membrane-targeted GFP. *Biotechniques* **24**, 349–350 (1998).
19. Blandino, G., Levine, A. J. & Oren, M. Mutant p53 gain of function: differential effects of different p53 mutants on resistance of cultured cells to chemotherapy. *Oncogene* **18**, 477–485 (1999).
20. Honigswachs, J., Faktor, O., Dikstein, R., Shaul, Y. & Laub, O. Liver-specific expression of hepatitis B virus is determined by the combined action of the core gene promoter and the enhancer. *J. Virol.* **63**, 919–924 (1989).
21. Pear, W. S., Nolan, G. P., Scott, M. L. & Baltimore, D. Production of high-titer helper-free retroviruses by transient transfection. *Proc. Natl Acad. Sci. USA* **90**, 8392–8396 (1993).
22. Canman, C. E. & Lim, D. S. The role of ATM in DNA damage responses and cancer. *Oncogene* **17**, 3301–3308 (1998).
23. Vogt Sionov, R. et al. c-Abl neutralizes the inhibitory effects of Mdm2 on p53. *J. Biol. Chem.* **274**, 8371–8374 (1999).

Acknowledgements. We thank W. J. Kaelin for the p73 α plasmid; T. Kanda for the H2B-GFP plasmid; W. Jiang and T. Hunter for E-GFP; S. P. Goff for the Abi plasmids; and R. Bernards for allowing some of the work to be performed in his laboratory. This work was supported by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities (Y.S.), by the Ebner Family Biomedical Research Foundation at the Weizmann Institute of Science in Memory of Alfred and Dolfi Ebner (Y.S.), by the NCI (M.O.) and by Telethon (Italy) (G.B.).

Correspondence and requests for materials should be addressed to Y.S. (e-mail: Ivshaul@weizmann.weizmann.ac.il).

p73 is regulated by tyrosine kinase c-Abl in the apoptotic response to DNA damage

Zhi-Min Yuan^{*†}, Hisashi Shioya[†], Takatoshi Ishiko[†], Xiangao Sun[†], Jijie Gu^{*†}, YinYin Huang[†], Hua Lu[‡], Surender Kharbanda[†], Ralph Weichselbaum[§] & Donald Kufe^{*†}

^{*} Department of Cancer Cell Biology, Harvard School of Public Health,

[†] Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA

[‡] Department of Biochemistry and Molecular Biology, Oregon Health Science University, Portland, Oregon 97201, USA

[§] Department of Radiation and Cellular Oncology, University of Chicago, Chicago, Illinois 60637, USA

The protein p73 is a structural and functional homologue of the p53 tumour-suppressor protein but, unlike p53, it is not induced in response to DNA damage^{1,2}. The tyrosine kinase c-Abl is activated by certain DNA-damaging agents³ and contributes to the induction of programmed cell death (apoptosis) by p53-dependent and p53-independent mechanisms⁴. Here we show that c-Abl binds to p73 in cells, interacting through its SH3 domain with the carboxy-terminal homo-oligomerization domain of p73. c-Abl phosphorylates p73 on a tyrosine residue at position 99 both *in vitro* and in cells that have been exposed to ionizing radiation. Our results show that c-Abl stimulates p73-mediated transactivation and apoptosis. This regulation of p73 by c-Abl in response to DNA damage is also demonstrated by a failure of ionizing-radiation-induced apoptosis after disruption of the c-Abl-p73 interaction. These findings show that p73 is regulated by a c-Abl-dependent mechanism and that p73 participates in the apoptotic response to DNA damage.

Because c-Abl mediates DNA-damage-induced apoptosis in a p53-independent way⁴ and p73 contributes to the induction of apoptosis^{1,2}, we investigated whether c-Abl interacts with p73 by immunoblotting anti-c-Abl immunoprecipitates from COS7 cells with anti-p73. The results demonstrate constitutive binding of c-Abl and p73; as a control, there was no detectable p73 in anti-RPA70 immunoprecipitates (Fig. 1a). As c-Abl is activated by DNA damage^{3,5-8}, we tested whether c-Abl associates with p73 after exposure to ionizing radiation. Analysis of anti-c-Abl immunoprecipitates with anti-p73 indicated that c-Abl bound to p73 in both control and irradiated cells (Fig. 1b). We also prepared ³⁵S-labelled p73 α and p73 β by *in vitro* transcription/translation for incubation with Abl fusion proteins with glutathione-S-transferase (GST). Analysis of the adsorbates demonstrated binding between GST-Abl SH3 and the p73 α isoform¹, but not with a GST fusion protein prepared from the c-Abl SH2 domain; results were similar with GST-Abl SH3 and p73 β (Fig. 1c). p73 contains an amino-terminal acidic transactivation domain (TAD), a central DNA-binding core domain (DBD) and a C-terminal homo-oligomerization domain (OD)¹. We incubated *in vitro* transcribed and translated proteins containing the different p73 domains with GST-Abl SH3. Analysis of these adsorbates demonstrated that the c-Abl SH3 domain binds to the OD, but not the TAD or DBD, of p73 (Fig. 1d). To confirm these findings, we assayed mutants of p73 β deleted at amino acids 311-419 and 419-494 for binding to c-Abl SH3. The results demonstrate binding of c-Abl SH3 to p73 β (Δ 419-494) but not to p73 β (Δ 311-419) (not shown), indicating that c-Abl associates with the p73 C-terminal homo-oligomerization domain.

To determine whether c-Abl phosphorylates p73, we incubated recombinant c-Abl with immuno-purified Flag-tagged p73 α and p73 β in the presence of [³²P]ATP. Analysis of the products by

autoradiography showed that both p73 α and p73 β are substrates for c-Abl *in vitro* (Fig. 2a, left), but there was no detectable phosphorylation of p73 α or p73 β when c-Abl was substituted with recombinant kinase-inactive c-Abl(K-R)⁹ (data not shown). In addition, c-Abl-mediated phosphorylation of p73 β was prevented in a p73 β (Y-F) mutant in which Tyr 99 (YAQP) was altered to phenylalanine (FAQP) (Fig. 2a, right). As a control, mutation of Tyr 121 (YGP) to phenylalanine had little if any effect on c-Abl-mediated phosphorylation (Fig. 2a, right). To test whether p73 is phosphorylated by c-Abl *in vivo*, we co-transfected 293 cells with c-Abl and GFP-p73 β , in which p73 β was tagged with green fluorescent protein (GFP). Analysis of anti-GFP immunoprecipitates with anti-phosphotyrosine antibody (anti-P-Tyr) demonstrated tyrosine phosphorylation of p73 β (Fig. 2b, left). By contrast, co-transfection of the kinase-inactive c-Abl(K-R) mutant and GFP-p73 β resulted in no apparent tyrosine phosphorylation of p73 β (Fig. 2b, left). In addition, there was no detectable c-Abl-mediated phosphorylation of the p73 β (Y⁹⁹-F) mutant (Fig. 2b, right). To assess whether p73 is phosphorylated in response to DNA damage, we irradiated COS7 cells which express high levels of endogenous p73 α and p73 β (ref. 10). Analysis of anti-p73 immunoprecipitates with anti-P-Tyr demonstrated tyrosine phosphorylation of p73 in irradiated, but not control, cells (Fig. 2c). To determine whether irradiation-induced tyrosine phosphorylation of p73 is c-Abl-dependent, we used MCF-7 cells which stably express the empty pSR vector or c-Abl(K-R)¹¹; these cells were transfected to express

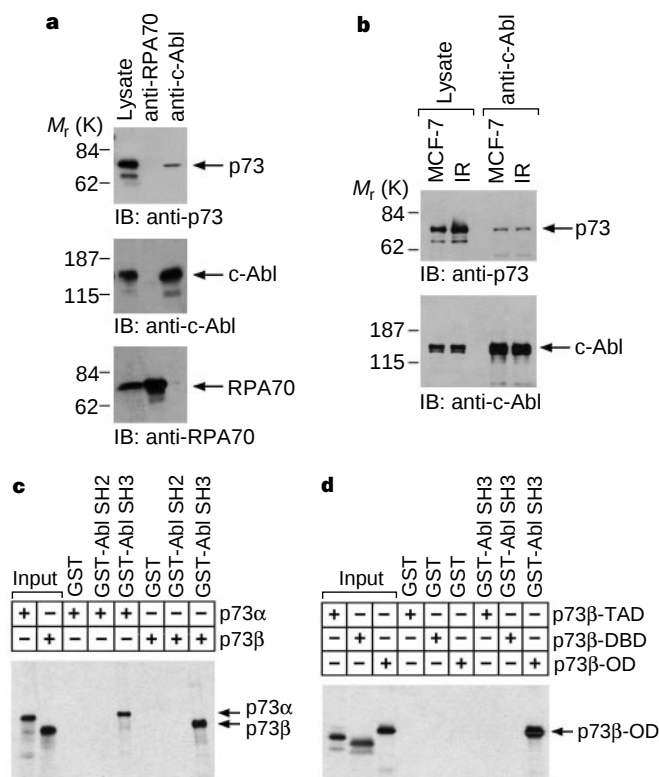


Figure 1 Association of c-Abl and p73. **a**, Lysates from COS7 cells were immunoprecipitated with anti-RPA70 or anti-c-Abl. The immunoprecipitates were analysed by immunoblotting (IB) with anti-p73, anti-c-Abl or anti-RPA70. Lysate was directly analysed by immunoblotting as a control. **b**, MCF-7 cells were exposed to 20 Gy of ionizing radiation (IR) and collected 1 h later. Lysates were analysed by immunoblotting with anti-p73 or anti-c-Abl (left two lanes), or immunoprecipitated with anti-c-Abl and the precipitates immunoblotted with anti-GFP, anti-p73 or anti-c-Abl (two right lanes). **c**, ³⁵S-labelled Flag-p73 α and Flag-p73 β were incubated with GST, GST-AblSH2 or GST-AblSH3 fusion proteins. Adsorbates were analysed by SDS-PAGE and autoradiography. **d**, ³⁵S-labelled Flag-p73 β TAD, Flag-p73 β DBD and Flag-p73 β OD were incubated with GST or GST-AblSH3. Adsorbates were analysed as in **c**.

GFP-p73 β and then exposed to 20 Gy ionizing radiation. Lysates were immunoprecipitated with anti-GFP and the precipitates were immunoblotted with anti-P-Tyr. The results demonstrate tyrosine phosphorylation of p73 β in irradiated, but not control, MCF-7/pSR cells (Fig. 2d). By contrast, there was no detectable tyrosine phosphorylation of p73 β in irradiated MCF-7/c-Abl(K-R) cells (Fig. 2d). To confirm that p73 is phosphorylated by a c-Abl-dependent mechanism, wild-type (Abl^{+/+}) and c-Abl-deficient (Abl^{-/-}) mouse fibroblasts¹² were transfected to express GFP-p73 β and then exposed to ionizing radiation: tyrosine phosphorylation occurred on p73 β in irradiated Abl^{+/+} but not Abl^{-/-} cells (Fig. 2e); there was no detectable phosphorylation of p73 β (Y⁹⁹-F) in irradiated Abl^{+/+} cells (data not shown). Radiation-induced c-Abl kinase activity is diminished in cells from patients with ataxia telangiectasia (AT)^{7,8}. To determine whether tyrosine phosphorylation of p73 is affected in AT cells, we transfected human diploid (HDF) and ATM-deficient⁷ AT5 fibroblasts with GFP-p73 β . Analysis of anti-GFP immunoprecipitates with anti-P-Tyr indicated that radiation-induced tyrosine phosphorylation of p73 β is diminished in AT5 cells relative to control cells (Fig. 2f). Taken together, these results show that c-Abl phosphorylates p73 on Tyr 99 in response to DNA damage.

To determine whether c-Abl affects p73 function, we transfected SAOS2 cells, which are deficient in both p53 (ref. 13) and p73 (ref. 1), with a construct containing the luciferase gene driven by a p53

enhancer from the *MDM2* promoter (mdm2NA-Luc)¹⁴ that had been used previously for p73 transactivation assays² (Fig. 3a). Co-transfection of mdm2NA-Luc with vectors expressing wild-type c-Abl or kinase-inactive c-Abl(K-R) had no detectable effect on luciferase activity (Fig. 3a). By contrast, co-transfection of mdm2NA-Luc and the GFP-p73 β vector resulted in activation of the luciferase reporter (Fig. 3a). Expression of c-Abl, but not of c-Abl(K-R), enhanced the transactivation function of p73 β (Fig. 3a). To assess this role of c-Abl in p73-mediated transactivation, we assayed SAOS2 cell transfectants for induction of p21. As shown previously¹⁵, expression of p73 β was associated with increased expression of p21 protein (Fig. 3b). Co-transfection of p73 β and c-Abl, but not c-Abl(K-R), strongly induced p21 compared with cells transfected with p73 β alone (Fig. 3b). As controls, transfection of c-Abl or c-Abl(K-R) in the absence of p73 had little, if any, effect on p21 expression (Fig. 3b). The finding that transfection of p73 β (Y⁹⁹-F) caused induction of p21 indicated that this mutant has an intact transactivation function (Fig. 3b). However, in contrast to wild-type p73 β , co-transfection of c-Abl with p73 β (Y⁹⁹-F) had no detectable effect on the induction of p21 compared with that mediated by p73 β (Y⁹⁹-F) alone (Fig. 3b). These findings show that c-Abl induces the transactivation function of p73 by phosphorylation of Tyr 99.

As transfection of SAOS2 cells with p73 results in the induction of apoptosis², we tested whether interaction between c-Abl and p73

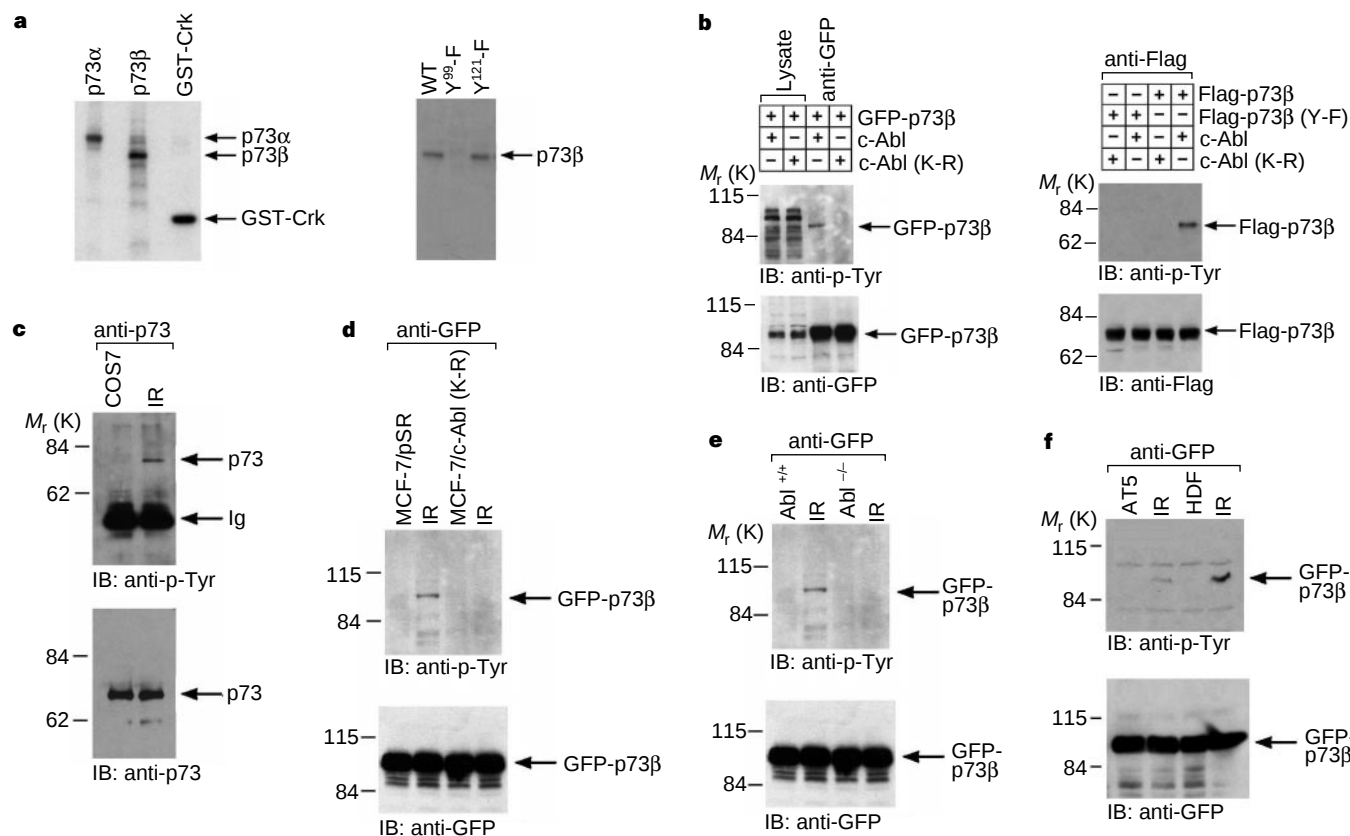


Figure 2 Phosphorylation of p73 by c-Abl. **a**, 293 cells were transfected with Flag-p73 α , Flag-p73 β , Flag-p73 β (Y⁹⁹-F) or Flag-p73 β (Y¹²¹-F). Immunoprecipitated proteins were incubated with recombinant c-Abl and [γ -³²P]ATP. GST-Crk(120–225) was used as a control. Reaction products were analysed by SDS-PAGE and autoradiography. **b**, 293 cells were transfected with GFP-p73 β and c-Abl or c-Abl(K-R) (left). After 48 h, lysates were immunoprecipitated with anti-GFP, and the precipitates immunoblotted with anti-P-Tyr and anti-GFP (right two lanes). Lysates were analysed by immunoblotting (left two lanes). 293 cells were transfected with Flag-p73 β , Flag-p73 β (Y⁹⁹-F), c-Abl or c-Abl(K-R) (right). After 48 h, lysates were immunoprecipitated with anti-Flag and the precipitates immunoblotted with anti-P-Tyr and anti-Flag. **c**, COS7 cells were exposed to

20 Gy of ionizing radiation and harvested at 1 h. Anti-p73 immunoprecipitates were immunoblotted with anti-P-Tyr and anti-p73. **d**, MCF-7/pSR and MCF-7/c-Abl(K-R) cells were transfected with GFP-p73 β . At 48 h, cells were exposed to 20 Gy ionizing radiation and collected after 1 h for preparation of lysates. Anti-GFP immunoprecipitates were immunoblotted with anti-P-Tyr and anti-GFP. **e**, Wild-type (Abl^{+/+}) and Abl^{-/-} fibroblasts were transfected with GFP-p73 β . At 48 h, cells were exposed to 20 Gy IR and harvested at 1 h. Anti-GFP immunoprecipitates were immunoblotted with anti-P-Tyr and anti-GFP. **f**, Human diploid fibroblasts (HDF) and AT5 fibroblasts were transfected with GFP-p73 β , incubated for 48 h, exposed to 20 Gy IR and then harvested at 1 h. Anti-GFP immunoprecipitates were analysed with anti-P-Tyr and anti-GFP.

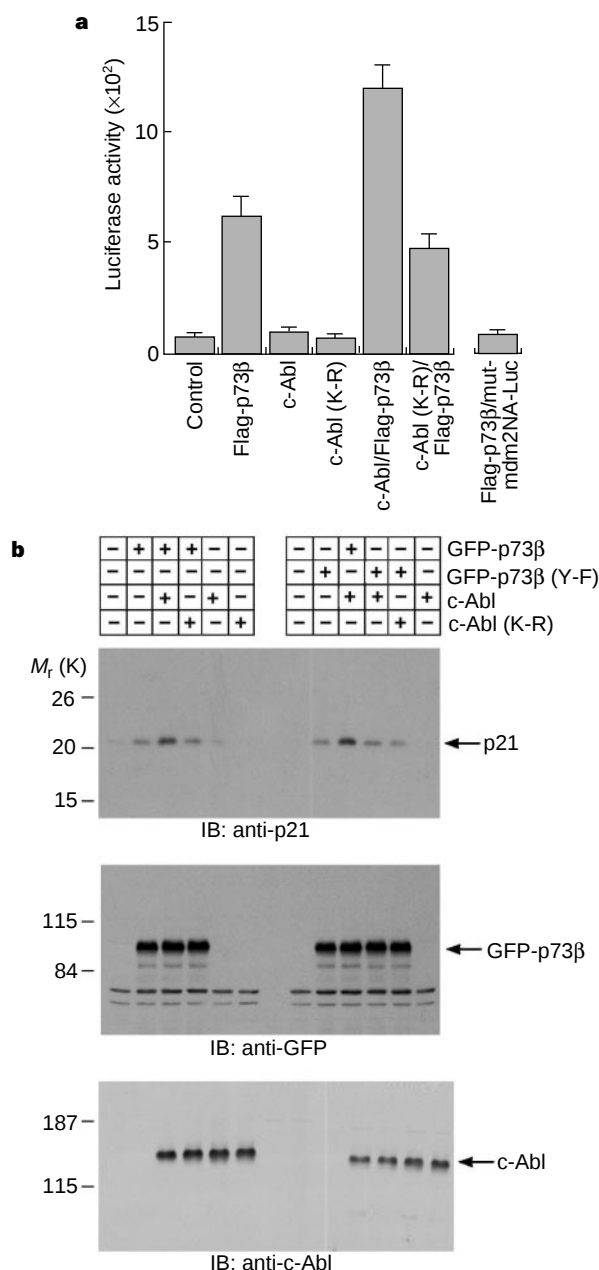


Figure 3 c-Abl induces p73-mediated transactivation. **a**, SAOS2 cells were co-transfected with 0.5 μ g p53-enhancer-luciferase plasmid (mdm2NA-Luc) and: (bar 1) 2 μ g Flag-pcDNA3 and 2 μ g pSR empty vectors; (bar 2) 2 μ g Flag-p73 β and 2 μ g pSR; (bar 3) 2 μ g Flag-pcDNA3 and 2 μ g pSR-c-Abl; (bar 4) 2 μ g Flag-pcDNA3 and 2 μ g pSR-c-Abl(K-R); (bar 5) 2 μ g Flag-p73 β and 2 μ g c-Abl; (bar 6) 2 μ g Flag-p73 β and 2 μ g c-Abl(K-R). SAOS2 cells were also co-transfected with 1 μ g p53-enhancer-luciferase plasmid containing a mutated p53 binding site (mut-mdm2NA-Luc), 2 μ g Flag-p73 β and 2 μ g pSR (right bar). Luciferase activity was measured 24 h after transfection and normalized for protein concentration to that for control vector. **b**, SAOS2 cells were transfected with the indicated vectors. Cells were collected at 24 h and their lysates immunoblotted with anti-p21, anti-c-Abl and anti-GFP.

affects p73-induced apoptosis. Whereas overexpression of c-Abl in MCF-7 cells and mouse embryo fibroblasts causes apoptosis⁴, transfection of the c-Abl vector into SAOS2 cells gave no apoptotic response (Fig. 4a); however, SAOS2 cells transfected to express GFP-p73 β showed an increase in the number of cells with sub-G1 DNA content compared with cells expressing empty vector (Fig. 4a). Co-transfection of GFP-p73 β and c-Abl(K-R) had little, if any, effect compared with the proportion of apoptotic cells resulting

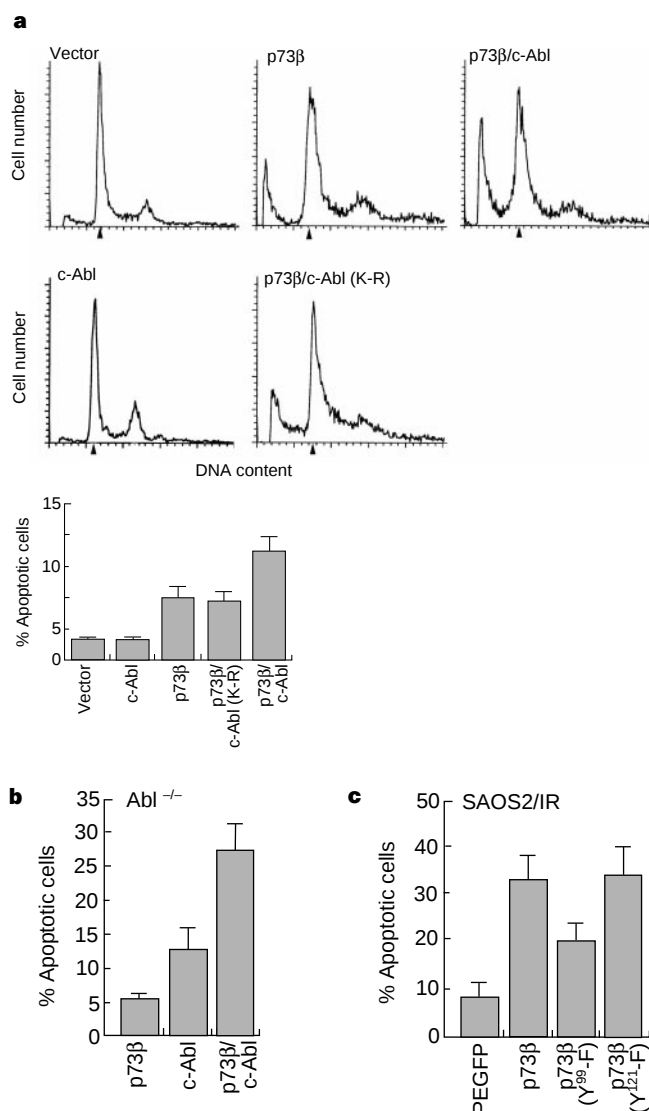


Figure 4 Interaction between c-Abl and p73 contributes to ionizing-radiation-induced apoptosis. **a**, SAOS2 cells were transfected with GFP-p73 β , GFP-p73 β (Y⁹⁹-F), c-Abl or c-Abl(K-R). Cells were collected at 48 h and assayed for DNA content (top panels). pEGFP was co-transfected with c-Abl. The results (mean $\pm$ s.e. of three experiments) are expressed as the percentage of cells with sub-G1 DNA (bottom panels). **b**, Abl^{-/-} cells were transfected with GFP-p73 β , c-Abl or with GFP-p73 β and c-Abl. pEGFP was co-transfected with c-Abl. The results (mean $\pm$ s.e. of two experiments, each done in duplicate) are expressed as the percentage of cells with sub-G1 DNA. **c**, SAOS2 cells were transfected with GFP, GFP-p73 β , GFP-p73 β (Y⁹⁹-F) or GFP-p73 β (Y¹²¹-F). At 24 h, cells were exposed to 20 Gy IR, collected after an additional 24 h and their DNA content assayed. Results (mean $\pm$ s.e. of two experiments, each done in duplicate) are expressed as the percentage of cells with sub-G1 DNA. The percentage of apoptotic cells expressing p73 β (Y⁹⁹-F), compared to p73 β or p73 β (Y¹²¹-F), is significantly different ($P < 0.05$).

from expression of GFP-p73 β alone (Fig. 4a). However, in contrast to the result with GFP-p73 β , co-expression of GFP-p73 β and c-Abl caused an increase in the number of apoptotic cells (Fig. 4a). To confirm the importance of the interaction between c-Abl and p73 in the induction of apoptosis, we co-transfected Abl^{-/-} cells with GFP-p73 β and either empty vector or c-Abl: we found that overexpression of GFP-p73 β had little effect on apoptosis of Abl^{-/-} cells (Fig. 4b). Co-transfection of GFP-p73 β and c-Abl, however, gave a substantial increase in the percentage of cells with sub-G1 DNA content compared with transfection with c-Abl alone (Fig. 4b). The

finding that c-Abl induces the apoptotic function of p73 suggests that p73 may be an effector of a c-Abl-dependent apoptotic response to DNA damage. To determine whether p73 is involved in apoptosis induced by ionizing radiation, we transfected SAOS2 cells to make them overexpress GFP-p73 β , p73 β (Y⁹⁹-F) or p73 β (Y¹²¹-F), and then irradiated them. There was little apoptosis in irradiated cells expressing the empty vector, whereas cells expressing p73 β or p73 β (Y¹²¹-F) showed an increase in the number of cells with a sub-G1 DNA content (Fig. 4c). By contrast, overexpression of p73 β (Y⁹⁹-F) blocked in part the apoptotic response to ionizing radiation (Fig. 4c). These findings collectively support a model in which c-Abl-mediated phosphorylation of p73 contributes to DNA-damage-induced apoptosis.

The apoptotic response of irradiated cells has been associated with expression of wild-type p53 (refs 16–18). c-Abl interacts with p53 in irradiated cells, but does not phosphorylate it, and contributes to radiation-induced G1 arrest by a p53-dependent mechanism¹¹. In contrast, c-Abl regulates DNA-damage-induced apoptosis predominantly by a p53-independent mechanism⁴. Our results support a model in which c-Abl regulates the p53-related p73 protein to induce DNA-damage-mediated apoptosis. Although it is known that p53 protein accumulates as a result of genotoxic stress¹⁹ and p73 does not^{1,2}, our results show that p73 is phosphorylated by a c-Abl-dependent mechanism in the DNA-damage response. The functional significance of the c-Abl/p73 interaction is supported by our findings that active c-Abl, and not inactive c-Abl(K-R), induces both p73-mediated transactivation and apoptosis. From our demonstration that preventing interaction between c-Abl and p73 also prevents radiation-induced apoptosis, we conclude that p73 is regulated by c-Abl in the DNA-damage response. We have provided evidence that p73 is activated by c-Abl kinase and that it participates in the apoptotic response to DNA damage. □

Methods

Cell culture. COS7, 293, MCF-7/pSR and MCF-7/c-Abl(K-R)⁴ cells were cultured in Dulbecco's modified Eagle's medium containing 10% heat-inactivated fetal bovine serum (HI-FBS), 2 mM L-glutamine, 10 units ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin. SAOS2 cells were grown in McCoy's 5a medium containing 10% HI-FBS, 2 mM L-glutamine, 10 units ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin. Wild-type (Abl^{+/+}), c-Abl-deficient (Abl^{-/-}), human diploid (GM00637F) and ATM-deficient AT5BIVA fibroblasts were all grown as described^{3,7,20}. Vectors were introduced into cells by using the Effectene transfection kit (Qiagen). Cells were treated with ionizing radiation at room temperature with a Gammacell 1000 (Atomic Energy of Canada) and a ¹³⁷Cs source emitting at a fixed rate of 0.21 Gy min⁻¹.

Plasmid construction. Vectors expressing GFP-p73 α and GFP-p73 β were generated by subcloning human p73 α or p73 β cDNAs¹ into pEGFP (Clontech). Vectors expressing Flag-p73 α , Flag-p73 β , Flag-p73 β TAD (amino acids 1–125), Flag-p73 β DBD (amino acids 128–313) and Flag-p73 β OD (amino acids 311–499) were generated by subcloning full-length PCR-generated products from the p73 α and p73 β cDNAs¹ into Flag-tagged pcDNA3. Vectors expressing Flag-p73 β (Y⁹⁹-F) and Flag-p73 β (Y¹²¹-F) were prepared by site-directed mutagenesis.

Immunoprecipitation and immunoblot analysis. Cell lysates were prepared in lysis buffer containing 0.5% Nonidet P-40 and immunoprecipitated as described¹¹ with anti-c-Abl (Ab-3; Oncogene Science), anti-RPA70 (Ab-1; Oncogene Science), anti-p73 (rabbit antiserum against p73; amino acids SAATPNLGPVPGML) or anti-GFP (Clontech). Proteins were separated on SDS–polyacrylamide gels and western-blotted with anti-p73, anti-c-Abl, anti-RPA70, anti-P-Tyr (4G10; Upstate Biotechnology Inc.) or anti-p21 (Abl-1, Oncogene Science) antibodies.

Fusion-protein binding assays. Purified GST, GST–Abl SH3 (ref. 21) and GST–Abl SH2 (amino acids 115–213) proteins (5 μ g) were incubated with ³⁵S-labelled Flag-p73 α or Flag-p73 β . GST–Abl SH3 was incubated with ³⁵S-labelled Flag-p73 β TAD, Flag-p73 β DBD or Flag-p73 β OD and the adsorbates analysed by SDS–PAGE and autoradiography.

In vitro kinase assays. Kinase-active c-Abl or kinase-inactive c-Abl(K-R)

purified from baculovirus was incubated with immunoaffinity-purified Flag-p73 α , Flag-p73 β , Flag-p73 β (Y⁹⁹-F), Flag-p73 β (Y¹²¹-F) or a GST–Crk control, and [γ -³²P]ATP in kinase buffer (20 mM HEPES, pH 7.5, 1 mM dithiothreitol, 10 mM MgCl₂) for 15 min at 30 °C. Phosphorylated proteins were separated by SDS–PAGE and analysed by autoradiography.

Luciferase assays. Transient transfections of SAOS2 cells with mdm2NA-Luc¹⁴ were done with the Effectene transfection kit and cells were collected 24 h after transfection. Luciferase was assayed with an enhanced luciferase assay kit (1800K, Analytical Luminescence).

Apoptosis assays. DNA content was assessed by staining ethanol-fixed cells with propidium iodide and monitoring by FACScan (Becton-Dickinson). Numbers of GFP-positive cells with sub-G1 DNA content were determined with a MODFIT LT program.

Received 6 May; accepted 19 May 1999.

- Kaghad, M. *et al.* Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. *Cell* **90**, 809–819 (1997).
- Jost, C. A., Marin, M. C. & Kaelin, W. G. Jr p73 is a human p53-related protein that can induce apoptosis. *Nature* **389**, 191–193 (1997).
- Kharbanda, S. *et al.* Activation of the c-Abl tyrosine kinase in the stress response to DNA-damaging agents. *Nature* **376**, 785–788 (1995).
- Yuan, Z.-M. *et al.* Regulation of DNA-damage-induced apoptosis by the c-Abl tyrosine kinase. *Proc. Natl Acad. Sci. USA* **94**, 1437–1440 (1997).
- Kharbanda, S. *et al.* c-Abl activation regulates induction of the SEK1/stress activated protein kinase pathway in the cellular response to 1- β -D-arabinofuranosylcytosine. *J. Biol. Chem.* **270**, 30278–30281 (1995).
- Kharbanda, S. *et al.* Functional interaction of DNA-PK and c-Abl in response to DNA damage. *Nature* **386**, 732–735 (1997).
- Shafman, T. *et al.* Interaction between ATM protein and c-Abl in response to DNA damage. *Nature* **387**, 520–523 (1997).
- Baskaran, R. *et al.* Ataxia telangiectasia mutant protein activates c-Abl tyrosine kinase in response to ionizing radiation. *Nature* **387**, 516–519 (1997).
- Sawyers, C. L., McLaughlin, J., Goga, A., Havlik, M. & Witte, O. The nuclear tyrosine kinase c-Abl negatively regulates cell growth. *Cell* **77**, 121–131 (1994).
- Marin, M. C. *et al.* Viral oncoproteins discriminate between p53 and the p53 homolog p73. *Mol. Cell. Biol.* **18**, 6316–6324 (1998).
- Yuan, Z. M. *et al.* Role for the c-Abl tyrosine kinase in the growth arrest response to DNA damage. *Nature* **382**, 272–274 (1996).
- Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. Neonatal lethality and lymphopenia in mice with a homozygous disruption of the c-abl proto-oncogene. *Cell* **65**, 1153–1163 (1991).
- Diller, L. *et al.* p53 functions as a cell cycle control protein in osteosarcomas. *Mol. Cell. Biol.* **10**, 5772–5781 (1990).
- Barak, Y., Gottlieb, E., Juven-Gershon, T. & Oren, M. Regulation of mdm2 expression by p53: alternative promoters produce transcripts with nonidentical translation potential. *Genes Dev.* **8**, 1739–1749 (1994).
- Zhu, J., Jiang, J., Zhou, W. & Chen, X. The potential tumor suppressor p73 differentially regulates cellular p53 target genes. *Cancer Res.* **58**, 5061–5065 (1998).
- Clarke, A. R. *et al.* Thymocyte apoptosis induced by p53-dependent and independent pathways. *Nature* **362**, 849–852 (1993).
- Lowe, S. W., Schmitt, E. M., Smith, S. W., Osborne, B. A. & Jacks, T. p53 is required for radiation-induced apoptosis in mouse thymocytes. *Nature* **362**, 847–849 (1993).
- Lowe, S. W., Ruley, H. E., Jacks, T. & Housman, D. E. p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. *Cell* **74**, 957–967 (1993).
- Kastan, M. B., Onyekwure, O., Sidransky, D., Vogelstein, B. & Craig, R. W. Participation of p53 protein in the cellular response to DNA damage. *Cancer Res.* **51**, 6304–6311 (1991).
- Luo, C. M. *et al.* High frequency and error-prone DNA recombination in ataxia telangiectasia cell lines. *J. Biol. Chem.* **271**, 4497–4503 (1996).
- Ren, R., Ye, Z.-S. & Baltimore, D. Abl protein-tyrosine kinase selects the Crk adapter as a substrate using SH3-binding sites. *Genes Dev.* **8**, 783–795 (1994).

Correspondence and requests for materials should be addressed to D.K. (e-mail: donald_kufe@dfci.harvard.edu).

correction

p73 is a human p53-related protein that can induce apoptosis

Christine A. Jost, Maria C. Marin & William G. Kaelin Jr

Nature **389**, 191–194 (1999)

It has come to *Nature's* attention that the title of this Letter is misleading. In fact, the p73 cDNA sequence used in this work was of simian, not human, origin.

The GenBank database accession number for this African green monkey p73 sequence is Y11419. □

Industry beckons Benelux's brightest

The best career prospects for scientists in the Benelux countries are in industry or administration. But the outlook in universities is improving too, with more postdoc jobs in Belgium, and many Dutch professors nearing retirement.

Quirin Schiermeier

The Benelux countries — Belgium, the Netherlands and Luxembourg — lie at the heart of the economic 'crescent' of Europe, an industrial ribbon extending from Milan in northern Italy through Germany to London. But Belgium and the Netherlands are not in the same league scientifically and, given their different cultural and political backgrounds, they cannot be viewed as a unit. The Netherlands is number one for science among the world's smaller countries, ranking ninth in the world and fourth in Europe.

Some 80,000 papers were published by scientists in the Netherlands between 1992 and 1996, according to the Institute for Scientific Information in Philadelphia, and the country ranks third in the world in terms of scientific output (number of citations).

State-funded research is conducted at 13 research universities and at seven research institutes run by the Netherlands Organization for Scientific Research (NWO), the country's largest sponsor of basic research. The Royal Netherlands Academy of Arts and Sciences runs seven life-science institutes, and the Dutch applied research organization, the TNO, has 14 institutes.

Scientific research is concentrated in the densely populated area between Amsterdam and Rotterdam, focusing on chemistry, biomedicine, astrophysics, materials and Earth

sciences. Scientific hot spots exist also in Groningen (protein engineering and molecular genetics) and in Eindhoven, where physics, electronic engineering and applied mathematics have a strong tradition. And the 'Food Valley' in Wageningen is one of the strongest areas in Europe for nutrition research and plant breeding.

Competition for academic jobs is intense as permanent positions in universities are rare. And a decline in total research spending since 1987, plus a shift in political priorities from research to primary education, have limited the financial scope of state-financed research (see *Nature* 394, 405; 1998).

"In the short term, we expect competition for limited academic jobs to remain fierce," says Walter Van den Elsen, employment adviser at the Association of Dutch Universities. There are 12,000 scientists at the Technical Universities in Delft, Eindhoven and Enschede, and at other faculties, and around 1,200 vacancies can be expected per year.

Vacancies at the Big Five

Less than a third of the 3,000–4,000 Dutch PhD students will pursue an academic career. Those who leave the alma mater can expect few problems: "With a PhD it is more or less easy to find a well-paid job in industry or administration," says Jan Van Den Boon, former head of the NWO's personnel department. Five multinational companies, Philips, Shell, Akzo, Unilever and DSM, account for most industrial career opportunities (see Table 1 in Supplementary information).

More young biologists stay in basic research than do young chemists, 80 per cent of whom move to industry after their PhDs. Physics graduates are "very rapidly absorbed by the labour market", says Van Den Boon. Astronomy students, for example, have good career prospects in industry and government departments. Research is coordinated by the Netherlands Research School for Astronomy.

According to a recent study on the career prospects of postdoctoral researchers, commissioned by Dutch employers and trade unions, 85 per cent of postdocs aim for lifetime careers in a university, but only a few make it. The study identifies the main problems as the need to 'job hop' between temporary posts, and the risk of being considered 'too old' for a position outside a university. The issue is to be addressed in a government report on science funding, due for publication this month. Despite long-term career uncertainties, the study found Dutch postdocs to be happy with their work and highly motivated, consistent with the survey carried out last year

by *Nature* and the European Science Foundation (see *Nature* 397, 640–641; 1999).

Young academic scientists have some reason to be confident. A considerable number of university professors will retire within five to ten years, taking a load off the strained academic job market, according to Van den Elsen. But by the time the new jobs appear there may be insufficient young scientists left to fill them. Science is going out of fashion in comparison with humanities, economics and law. "The Netherlands will soon face a lack of scientists," says Van den Elsen. Some universities and research institutes have already started to recruit foreign scientists.

The Association of Dutch Universities has been promoting the image of science in schools, and the Ministry of Science and Education has set up programmes to maintain the high standards of university research. In 1997, it launched the Van der Leeuw Programme, intended to provide a reasonable career perspective for talented researchers. The DFL80 million (US\$39 million) scheme makes it possible to appoint highly qualified successors to older professors some years before they retire.

In addition, each year five to ten top-level scientists below the age of 40 are awarded one of the desirable PIONIER grants. Generous funding (DFL200,000–400,000 a year) allows grant holders to set up their own teams and laboratories, and the host university guarantees the continuation of the research after the five-year subsidy period.

Belgium's divided cultures

In contrast to the Netherlands' highly successful scientific base, Belgium is a below-average research performer. According to the Institute for Scientific Information, 38,000 papers were published in Belgium between 1992 and 1996, ranking the country only ninth in Europe, and fourteenth in the world. Ranked by citations per paper, however, Belgium comes a respectable seventh, ahead of Canada, Germany, France and Italy.

Belgium is handicapped by its history. It is divided politically into three regions — Wallonia in the south, Flanders in the north and Brussels — and culturally into three communities, French and Dutch speakers, and a German-speaking minority. This divides research funds and cultures to the country's detriment. The Flemish Fund for Scientific Research (FWO) is strictly separated from its Wallonian counterpart (FNRS), and the resulting small scientific communities cannot be effective political lobbies.

Science funding is determined at federal,

Search the web for jobs

Netherlands Organization for Scientific Research

<http://www.nwo.nl>

Dutch Ministry of Education, Science and Culture

<http://minocw.nl>

Netherlands Research School for Astronomy

<http://www.strw.leidenuniv.nl/nova>

Dutch Applied Research Organization

<http://www.tno.nl>

Job sites

<http://www.AcademicTransfer.nl>

<http://www.jobnews.nl>

<http://www.monsterboard.nl>

Brussels Technopole

<http://www.technopol.be>

Flemish Institute of Biotechnology

<http://www.vib.be>

Belgian Bioindustries Association

<http://www.bba-bio.be>

Free University of Brussels

<http://www.vub.ac.be>

regional and community levels. Basic research at Belgium's ten universities is funded by the communities, application-orientated research by the regions, and the federal government pays for international collaborations (see Table 2 in Supplementary information).

When it comes to budgets, science has a low priority. In 1998, total public spending on science was only BEF59.4 billion (US\$1.6 billion), a mere 0.42 per cent of the aggregated public budgets. The situation is slowly changing for the better, but most jobs for scientists are provided by industry.

Some scientists complain that the little money available is distributed among too many universities. "We seldom reach the critical mass necessary for high-quality basic research," says Pierre Van Antwerpen of Technopole, a technology service institute in Brussels. "But merging or closing university departments is extremely difficult, for ideological reasons." Brussels, for example, has a French-speaking (ULB) and a Dutch-speaking university (VUB), both of which cover the whole range of sciences. There is also the Catholic University of Brussels and the Brussels site of the Catholic University of Leuven.

Science in Brussels lags behind other European capitals, but it has strengths in medical research and life sciences. There is excellent research in microbiology, immunology, parasitology and protein engineering.

Walloons wake up to opportunities

Wallonia has a relatively poor research base, with modest strengths in astrophysics, space technology, biotechnology, materials and information technology. The mainly rural region has high unemployment—a legacy of its collapsed coal mining and steel industries.

"A university research career is tough," says Etienne Reuter, an official at Wallonia's General Directorate for Technology, Research and Energy (DGTRE). But, since there are few qualified scientists in Wallonia, the job prospects in industry are relatively good, says Reuter. Jobs are available at international companies, such as Pfizer, SmithKline Beecham Biologicals, Baxter Healthcare, Solvay and UCB Pharma. And there are signs that the Wallonian government has recognized the role that science may play in reflatting the economy. Over the past five years, the regional budget for economically relevant research has almost doubled, to BEF5 billion. A modern science base is to be set up along the Brussels–Namur–Luxembourg axis.

The DGTRE supports spin-offs from the research universities in Louvain-la-Neuve, Liège and Mons, each of which has a well-established science park. It is up to the universities to choose which companies are admitted to the parks. More than 3,000 jobs have been created in the fast growing Louvain-la-Neuve Science Park, for example, and 25 per cent of its 88 companies are university spin-offs. One company, Ion Beam Applications,

Belgian biotech boom brings skill shortage

The biotechnology industries of Belgium and the Netherlands are important components of the scientific landscape in both countries. Biotechnology's relative importance for the job market, however, is bigger in Belgium.

There are around 60 biotechnology companies in the Netherlands, and around 50 in Belgium, according to a recent report by the consultancy Ernst & Young. The Netherlands and Belgium rank seventh and eighth in Europe in terms of the number of biotechnology companies, says the report.

In 1997, biotech companies in Belgium employed 4,500 people and invested BEF8.6 billion in research. Two of Europe's financially most successful biotech companies, Innogenetics and Plant Genetic Systems, are based in Ghent.

The regional Belgian governments provide increasing amounts of financial support for new initiatives. This year, the Flemish government approved a BEF110 million grant to expand Pharming, a Dutch pharmaceutical

company. The money will be used to build a plant in Geels. At least 50 scientific jobs will be created.

In 1995, the Flanders Interuniversity Institute for Biotechnology was set up near Ghent to coordinate in a single autonomous institute the work of nine top university research teams and five associated laboratories. One of the institute's goals is to found spin-off companies, such as DevGen and CropDesign. The institute, employing 700 scientists, is sponsored by the Flemish government to the tune of BEF947 million a year.

Career prospects in Belgium's biotechnology industry are excellent. "We have to employ senior scientists from abroad," says Désiré Collin of the University of Leuven. Flemish universities now offer a masters degree in biomedical science to meet the increasing demand for staff.

There is no analogous biotechnology boom in the Netherlands. "We are light-years behind developments in Britain and Germany," says Frank Grosveld, a molecular

biologist at the University of Rotterdam. Successful young companies, such as Leiden-based Introgene and Pharming, are the exception. During 1997 and 1998 only ten start-ups entered the market—far less than in Switzerland, Sweden or Belgium.

The lack of big pharmaceutical companies that could serve as catalysts for start-ups is viewed by many as the main problem. And the legal framework is very strict in the Netherlands, where public attitudes to biotechnology tend to be hostile. So some companies have started operations across the Belgian border. "There is also a lack of entrepreneurial culture at Dutch universities," says Menno Horning of the Ministry of Economic Affairs.

The Dutch government is starting a DFL30 million life-science programme this summer, aimed at increasing the number and quality of biotech companies. "We want to stimulate creative young scientists to start their own business rather than selling licences to big companies," says Horning. **Q.S.**

producing particle accelerators for medical and industrial applications, was founded in 1986 by Yves Jongen, a former student of the local university. Today, the company employs 220 people, half of them scientists.

Flanders has a stronger tradition in science. The University of Leuven conducts cutting-edge research, particularly in gene therapy. Other Flemish universities are at Ghent and Antwerp, the latter hosting the renowned Institute of Tropical Medicine. Applied environmental, energy and materials research is conducted at the Flemish Institute for Technical Research at Mol.

Flanders receives the largest share of Belgium's science budget (BEF24.7 billion). Over the past five years, university grants distributed by the FWO have nearly doubled, to BEF3.7 billion, and a further increase of BEF2 billion is promised. FWO-funded PhDs and postdoctoral fellowships have tripled to 220.

"Additional funding has given a serious impetus to science in Flanders," says José Traest, general secretary of FWO. The job outlook in Flemish industry is good. There are barely enough scientists and engineers to

meet the demands of companies such as Janssen Pharma, Pharming, and microelectronics and telecommunications companies. The job situation in the academic sector is slowly recovering. "Career opportunities have improved in the past two years, even though spending on science is still too low," says Désiré Collin, head of molecular and vascular biology at the University of Leuven.

Research experience outside Belgium is almost imperative for an academic career in Flanders. "Basic university education is quite good, but after that young academics are well-advised to go abroad for a few years," says Collin. The university system is notorious for its inflexibility. Many professors cling to their vested rights. Recently, however, quality control has been introduced: the performance of professors is to be evaluated. "Knowing that people who don't come up to scratch could be fired will surely change the atmosphere for the better," says Collin. □

Quirin Schiermeier is in Nature's Munich office.
e-mail: a.abbott@nature.com

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>).